DOCUMENT RESUME

ED 214 297 EA 014 560

AUTHOR Bank, Adrianne; Williams, Richard C.

TITLE Evaluation Design Project: School District Organization Study. Annual Report.

INSTITUTION California Univ., Los Angeles. Center for the Study of Evaluation.

SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
PUB DATE 1 Dec 80
GRANT NIE-G-80-0112-P-5
NOTE 247p.

EDRS PRICE MF01/PC10 Plus Postage.
DESCRIPTORS Case Studies; Classroom Techniques; Coordination; Educational Cooperation; Educational Environment; *Educational Testing; Elementary Secondary Education; *Instructional Improvement; Models; Questionnaires; School Districts; *Student Evaluation; Teacher Attitudes
IDENTIFIERS *Linkage Analysis; Naturalistic Research

ABSTRACT

A three-year research project aims to describe and analyze how school district management of educational testing and evaluation can better link these activities to instructional improvement, according to this first annual report. After chronicling the first year's activities, the authors discuss their use of case studies, interviews, and naturalistic inquiry to gather data on the district subsystems linking testing, evaluation, and instruction, the subsystems' origins and characteristics, and the effects of testing and evaluating on teachers and classroom activities. They review the literature on the subject and construct an analytical framework that includes social, professional, and district environments, educators' ideas about testing-evaluation-instruction (T-E-I) linkages, the components of T-E-I subsystems, staff performance, coordinating mechanisms, and classroom impacts. The report describes the four case studies (of a projected six) begun so far, presenting for each district historical and background information, district testing and professional development functions, linking activities and mechanisms, and classroom effects. Each case study is reexamined in terms of the analytical framework. Observations based on the case studies are then made about the framework's factors. The project questionnaire and supporting documents are appended. (RW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED214297

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Gabriella Gray

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

DELIVERABLE - December 1, 1980

EVALUATION DESIGN PROJECT: SCHOOL DISTRICT ORGANIZATION STUDY

Annual Report
December, 1980

Adrianne Bank and Richard C. Williams
Study Directors

Adrianne Bank and Richard C. Williams
Project Directors

Grant Number
NIE-G-80-0112
P-5

EA 014 560

CENTER FOR THE STUDY OF EVALUATION
Graduate School of Education
University of California . Los Angeles

TABLE OF CONTENTS

	Page
ACKNOWLEDGMENTS	
SUMMARY	
PART I:	
Purpose	2
Background	5
Chronology of activities	9
Issues in methodology	14
PART II:	
Perspectives from the literature on school district use of testing and evaluation for instructional improvement	22
Preliminary Analytic Framework of T/E/I linking subsystems	40
Case studies of four districts	49
Applications of Analytic Framework	131
Observations from four case study districts	146
PART III:	
References	
Appendices	

ACKNOWLEDGMENTS

As the authors of this document, we would like to thank the members of the staff of the Evaluation Design Project whose headwork, fieldwork and deskwork contributed so much: Alison Arcuri, Bruce Convissar, Don Dorr-Bremme, Ann Picker, Michele Marcus, Pamela McGranahan.

We would also like to thank others at CSE who discussed the issues or reviewed the language contained in this document: Eva Baker, Jim Burry, Richard Daillak, Joan Herman, Jennie Yeh.

Our Working Group must remain anonymous for a time. However, as they themselves know, their support, encouragement and experience were invaluable to us. So, too, were the insightful comments and suggestions of our consultants, Dr. Ann Lieberman, Associate Professor, Teachers College, Columbia University, and Dr. E. Mark Hanson, Associate Professor, University of California, Riverside.

Finally, we will never forget the patience and goodwill with which Aurelia Arroyo transformed our illegible scribbles into respectable typescript.

SUMMARY

This document is a draft of the December 1980 Annual Report of the Evaluation Design Project: Organizational Study being coordinated by the Center for the Study of Evaluation (CSE) at UCLA. It is being submitted to the National Institute of Education (NIE) for review concurrent with additional reviews occurring within CSE.

Within Part I, we briefly outline the purpose and background of the project and present a chronology of activities for the year 1979-1980. Following this we informally discuss our methodology in relation to three problem areas often encountered in naturalistic inquiry.

Within Part II, we review selected research and theoretical literatures that have helped to sharpen our focus and thereby improve our grasp on the phenomenon we are examining. The preliminary analytic framework which emerged from this literature review, from our interview data and from interaction with our Working Group is described and used to organize presentations of the Case Studies and the brief analytic descriptions which follow them. We conclude with some observations and issues.

Part III contains references and and appendices including our Questionnaire Modules.

PART I.

PURPOSE

The Evaluation Design project located at the Center for the Study of Evaluation (CSE) has completed the first year of its three-year (1979-1982) effort. The project represents one aspect of CSE's long-term plan to conduct research into the ways in which testing and evaluation activities may be designed and conducted so as to beneficially impact on instruction and student learning. Our intentions remain as stated in our original plan "... to examine existing district policies, procedures and programs that appear to contribute to systematic use of evaluation and testing for educational improvement and make recommendations based on that examination for district guidance." (8/30/79, p 6)

The primary purpose of the Evaluation Design project, then, is to describe and analyze the management of testing and evaluation activities in those districts which have indicated that it is their intention to link testing and evaluation with instructional improvement.

We begin with the following assumption: if tests and evaluations which do not originate with teachers are to influence teachers' thinking and change instructional activities in classrooms in any directed or sustained fashion, then the district office's system for supporting the use of test scores and of evaluation findings is a vital factor. We are therefore examining the management subsystems within districts that link testing and evaluation with instructional

improvement. We are also examining the effect that these subsystems have on teachers' thinking and on classroom activities.

It should be noted that management is one of several variables that can influence the use of evaluation and testing data. Other variables relate to the quality of available instruments (Boyd, et al, 1975), to the conflict between assumptions about learning made by test-makers and by teachers (Lortie, 1975), to teachers' lack of training in the design, interpretation and uses of tests (Cromer & Slakter, 1968). Our three years of research will not decisively test our assumption that district management is a most important factor in whether tests or evaluations are used in ways that lead to instructional improvement. Such a validation study is an important next step for future research. Before we can do that we must first be able to describe what districts' T/E/I linking subsystems can look like.

At the conclusion of our three-year research, we will have described such subsystems in six districts which have developed or are in the process of developing them. We will have analyzed the components of those subsystems and their impact on teachers and on classroom activities. We will be able to describe: how such subsystems came to be; what seem to be characteristics essential to their functioning; if and how they have affected teachers' attitudes and intentions; if and how classroom activities are impacted. At the conclusion of our first year, we have a preliminary version of such an analytic framework as well as observations from four districts. In our second and third years we will refine that framework by continuing research in two of the original districts and in two additional districts.

A second purpose of the Evaluation Design project is to produce self-help technical assistance materials for districts who wish to develop a management subsystem for using testing and evaluation to improve instruction. By the end of the third year, we will have developed and pilot tested several products related to establishing and operating such a subsystem as well as assessing its impact. We will also have incorporated into a plan our investigation of ways of disseminating and promoting the utilization of such products.

Products and findings from the Evaluation Design project will provide a number of benefits:*

- They will fulfill NIE's mandate to conduct research on issues related to the improvement of practice.
- They will support NIE's intention to develop school district capacity for integrating evaluation, testing and instruction.
- They will provide school district evaluation professionals with a sense of their own role in relation to educational renewal.
- They will provide school district evaluation professionals with a sense of the similarities and differences between their own policies, procedures and problems and those of other districts.
- They will make possible the development of empirically based strategies for providing technical assistance to school districts to improve the linkages among testing, evaluation and instruction.
- They will provide accurate information on the organization and technical issues constraining or facilitating the linking of evaluation, testing and instruction.

---

*Excerpted and paraphrased from Evaluation & Decision Making Program, proposal submitted to NIE for period June 1, 1978 through November 30, 1979 (pp 7-8).

BACKGROUND

Since 1965, CSE has taken as its mission the exploration of problems associated with testing and evaluation. During the late '60s and early '70s, our research was directed towards the generation of evaluation techniques and models. We incorporated these techniques and models into kits and workshops for training purposes. By the mid 70s, however, we had acquired a great deal of experience with the limitations of short-term evaluation training. This experience, which was shared by our colleagues at other institutions, led us to believe that research on the local context in which evaluation activities occur was essential; only after such research could we understand how evaluation and related testing programs might lead to school improvement.

So it was that in CSE's 1977 Long Range Plan we asserted our intention "to study and improve the utility of evaluation as it occurs within the educational setting." We went on: "... in doing so we will take the role of national consumer advocate in the area of evaluation. Our advocacy is in the interests of the consumers of educational services, the students and supporters of the educational enterprise, the professional and lay communities." (p 3)

Subsequently in our Resource Allocation and Management Plan of that year three broad lines of inquiry into local conditions were elaborated

---

*Parts of this section appeared originally in Bank, A., Interrelationships Among Test Use, Evaluation Use and Evaluation Design Projects, June 1980, and Williams (Ed.), Faculty Papers Dealing with School District Research, Development and Evaluation Units, November 1979.

--how teachers use tests in their classrooms, how program managers of specially-funded projects use evaluation information to make decisions, and how school districts organize their testing and evaluation operations.

The research into how school districts organize their testing and evaluation operations had two thrusts --one descriptive using a nation-wide survey, the other theoretical deriving from a sociological analysis of school district organizations and the technology of evaluation and testing.

The survey of districts having centralized offices responsible for testing and evaluation activities found the following (Lyon, et al, p 5-6):

- Formal school district evaluation is a relatively young and growing field in education. Fifteen percent of existing evaluation offices were established before 1965. Since 1970, the creation of these offices has accelerated, with 51% newly organized or reorganized from other units (testing, research) since then.

- School districts typically allocate a small percentage of their total budget to central evaluation offices. Ninety-seven percent report budgets of 1% or less of the district budget. Evaluation unit budgets range from $2,000 to $4,000,000, representing per pupil expenditures ranging from $.10 to $90. Half of the respondents report that their budgets are adequate.

- Two-thirds of the respondents reported that their personal resources are not adequate. The single most important problem reported is too few staff.

- Evaluation offices are involved in elementary, intermediate and secondary education. Approximately 75% agree that student achievement is the dominant topic of data collection. Approximately 75% agree that testing is their major method of data collection. Offices express little agreement on what constitutes basic evaluation practice or on the priority of various evaluation activities.

- Most evaluation offices administer testing programs. Seventy-one percent administer districtwide norm-referenced

testing (NRT) programs. Forty-eight percent administer districtwide criterion-referenced testing (CRT) programs.

The analyses, presented in the form of papers, were authored by David O'Shea, Charles O'Reilly, Oskar Grusky, and Lynne Zucker. O'Shea argues that school districts as institutional rather than technical organizations do not have the highly developed technical core relating to instruction which current evaluation studies presuppose. What is more, the nature of school districts is to emphasize form over substance, symbol over reality. Given these proclivities, it is to be expected that evaluation findings will have limited impact on school districts' technical, i.e., instructional, decisionmaking.

O'Reilly's very comprehensive review of the decisionmaking research literature reminds evaluators of the preeminence of the human factor in decisionmaking. Decisionmakers are not rational, consistent, and logical; they are not disposed to make decisions on the basis of "hard" data regardless of the consequences. Instead they usually consider many factors, only one of which is data, and they often arrive at decisions that reflect preconceptions or serve their own welfare. It is quite common, and completely understandable, for decisionmakers to use evaluations, or selected portions thereof, to confirm what they already had planned to do and/or to ignore disconfirming data.

Grusky combines additional data analysis of the survey questions with the literature on role conflict and ambiguity to point out those characteristics and behaviors of complex organizations that must be considered by evaluation unit directors who seek to have their unit's

work have decisionmaking impact. His analysis contains organizational, structural, and behavioral implications for those who design and lead school district evaluation units.

Finally, Zucker, through a further analysis of the survey data, illustrates how school districts utilize evaluations to perform a signaling function to the external environment. This signaling behavior, typical of institutionalized organizations responding to societal demands, is further stimulated because evaluations are usually targeted toward external funding agencies. As a result, she asserts that most evaluations stress ambiguity and a narrow range of information in order to serve an external function; their usefulness for informing internal decisionmaking is thereby diminished.

The 1979-82 work of the Evaluation Design project builds on this earlier work. Because of it, our current focus has broadened: from an emphasis on the formal organizational unit with the school district, called the research evaluation office, to the less defined construct called the T/E/I linking subsystem; from a concern primarily with evaluation and testing to a concern with the way in which all the operations of the district affect the relationships between testing, evaluation and instruction.

This project and the earlier work are linked conceptually. They are also linked by personnel. One of the current Evaluation Design Project Directors conceived the original plan for this work as it was presented in the 1977 Long Range Plan and the Resource and Allocation Management Plan. The other Director was part of the original survey staff and directed the project during its last six months.

CHRONOLOGY OF ACTIVITIES

In December 1979, we began work on the first year of the Evaluation Design Project in accordance with the plan proposed to the National Institute of Education (8/30/79). We intended to do fieldwork in four districts that had a reputation for having linked testing and evaluation with instructional improvement in some manner.

Our rationale for this work derived in part from previous CSE research on the use of testing by teachers (Yeh, 1978), the use of evaluations by project managers (Alkin, 1979), and especially from the survey of district research and evaluation units (Lyon, et al, 1978, 1979) and the analyses in CSE-commissioned papers written by a UCLA interdisciplinary faculty panel (Williams, 1979). These efforts supported the assumption that the use of evaluation or test data for the improvement of instruction --at the district, school and classroom levels-- was not a naturally occurring phenomenon. Rather, we believed that such use would only occur with purposeful and effective district effort. We further believed that most school districts had not formulated the system-supported procedures needed to connect evaluation or testing with instruction; and that, even in those that had, impact of district policies and practices on the thinking and behaviors of teachers might be problematic.

In view of the amounts of money spent on testing and evaluation, we wanted to explore the potential for district linkages among testing and evaluation by examining districts already moving in that direction. We wanted to learn the history of these districts' efforts, their current activities, and the impact on teachers and classroom events.

We started with no hypotheses about what we would find. We decided to collect the data around essentially descriptive questions; we would

then try to discover some patterns or framework from the data. These questions were outlined in our Progress Report (2/80, p 3):

1. WHY DO SCHOOL DISTRICTS INITIATE AND MAINTAIN SYSTEM-SUPPORTED LINKAGES AMONG TESTING/EVALUATION AND INSTRUCTIONAL ACTIVITIES?

2. HOW HAVE DISTRICTS' ACTIVITIES AND LINKAGES IN RELATION TO T/E/I CHANGED OVER TIME?

3. HOW DO SCHOOL DISTRICTS CURRENTLY MANAGE LINKAGES AMONG T/E/I?

4. WHAT IMPACT DO SYSTEM-SUPPORTED LINKAGES HAVE?

The following is a brief chronological account of our research activities from December 1, 1979 to October 1, 1980. Where appropriate we have commented briefly on problems we encountered and decisions that we made. Methodological issues will be treated in more detail in the next section.

Selected Sample Districts

Based on nominations from CSE colleagues, from members of AERA's Division H, other research colleagues, and selected school and government officials, and on results from previous CSE research (Lyons, et al, 1978) we identified 22 districts as possible research sites. Telephone inquiries and personal visits reduced the sample to eight geographically proximate districts. Final district selection was based on several criteria: district progress on linking evaluation and testing to instructional improvement, district approach or strategy and the districts' willingness to participate. As a set, we wanted the districts to reflect a variety of size, ethnic makeup, and socioeconomic characteristics. Project staff visited each of the four selected districts to finalize arrangements.

Formed Working Groups

Each participating district plus one other district which is associated with the Huron Institute's study, designated a staff member

who was associated with their linkage effort to serve on a Project Working Group. This Working Group served both as respondents to our research inquiries and as colleagues who helped us design, execute and evaluate our research study. This group met three times during the first year.

Developed Preliminary Research Design

Based on our original proposal to NIE (8/30/79), project staff developed an initial research strategy and preliminary interview protocols. The set of interview protocols was field tested in a nonparticipating district. Appropriate respondents in each of our districts were identified. The questionnaires were color coded to reflect the subject matter so that they could be modified if unexpected time constraints emerged. Research assistants were selected and all project staff attended field research training sessions. At those sessions, question ambiguities were clarified, the study's intent was explicated, and note taking and interviewing techniques were practiced.

Held First Working Group Meeting

The research design and the field-tested interview protocols were reviewed with the Working Group. Issues of entry, confidentiality, internal use of data by districts were discussed. Each district representative worked with a project staff member in planning field site visits to his or her district, including the scheduling of interviews and the securing of appropriate documents and reports.

Conducted First Field Visit

Two project staff members visited each district for two days collecting data on present district linking activities and strategies.

Each staff member visited two districts so that he/she could have multiple site experience. Interviews were held with: Superintendents, Board members, other administrators, evaluation, testing, curriculum and instructional specialists. Critical documents were obtained and examined. Approximately 15 interviews were held in each of the smaller districts; 20 interviews were held in the larger districts.

Held Second Working Group Meeting

Research progress was reviewed with Working Group members, including a summary of our findings in each district. Each Working Group member reviewed the data collected from his/her district for accuracy and reflected on our interpretations. Suggestions for thematic approaches and new questions raised by the data were offered. Discussions were held as to next steps to be taken. Schedules for the next visit were finalized. The Working Group suggested ways of measuring impact in classrooms and expressed their views that such information would be highly useful to them.

Conducted Second Field Visits

The same project staff members revisited each district for two days, focusing this time on the history of the District's efforts and the impact of the district-supported T/E/I linkage subsystem on teachers. We also filled in the gaps from our first visit and resolved conflicts and ambiguities in the data. Interviews were held with knowledgeable and long-time administrators, school Board members, community leaders, principals and teachers. Interviews concerning impact were held with from 20-40 teachers in each district.

Coded and Synthesized Data - Wrote Case Study of Each District

Files were established for each district and a strategy for data reduction and synthesis was developed. Interview notes were checked

against tape recordings of each interview. Each staff member drafted a case study and served as a reviewer of another case study. Staff meetings were held to distill out common themes and elements from the various case studies.

Developed Preliminary Analytical Framework

Based on between and within district analyses and the literature review, a preliminary analytical framework was developed and reviewed with an outside consultant. Appropriate revisions were made. Further details on this process are in the next section.

Held Third Working Group Meeting

Each district representative reviewed his/her district's case study for accuracy and interpretation, and read descriptions of the other districts. Our preliminary analytical framework was presented, discussed, and revised. Each representative then utilized the framework in analyzing his/her district's T/E/I linking subsystem. The project's second year's activities and dissemination strategies were also discussed.

Revised Framework and District Vignettes - Planned 2nd Year Activities

The preliminary analytical framework was revised, the second-year activities were determined, sequenced, budgeted, and preliminary dissemination strategies were outlined. These are described in a separate document.

ISSUES IN METHODOLOGY

During the first year of the Evaluation Design project, we conducted case studies in four school districts which were purported to be developing linkages between testing, evaluation and instruction. As much as possible we followed a naturalistic mode of inquiry. We will describe the methodological issues we faced using Guba's (1978) three categories which constitute "middle range problems" in naturalistic inquiry. These are: boundary problems, focusing problems and problems of authenticity.

Under our discussion of boundary problems, we will explain how we defined the study, selected the sample districts, and developed interview protocols.

Under our discussion of focusing problems, we will explain how we selected respondents within districts and how we used the interview protocols.

Under authenticity problems we will explain how we analyzed the data and interpreted the findings.

BOUNDARY PROBLEMS

Definition of the study. Guba notes that "Naturalistic inquiries are not mounted in vacuo; like all other forms of inquiry they emerge in response to a perceived problem" (p 44).

The problem as we perceived it, quite simply, was that school districts are being assaulted by external requirements for testing and evaluation. These requirements come from federal and state efforts to solve educational problems. Testing and evaluation in connection

with these efforts were often being used for monitoring purposes --that is, to ensure compliance-- or for information which, hopefully, would lead to policy revision. Whatever the virtues of these two intentions, district officials as well as teachers often perceive the testing and evaluation requirements as burdensome and irrelevant to their own efforts to deal with local problems. We became interested in whether externally-imposed requirements could be adapted or modified in some way by districts so as to meet local needs. From our own experience and from our knowledge of the literature, this seemed unlikely. Given what we know about the characteristics of school districts as complex organizations, and the role and function of teachers we thought that testing and evaluation might be always and immutably regarded as a foreign intrusion. On the other hand, our previous work at CSE had revealed an astonishing lack of attention, either from researchers or practitioners about the management of testing and evaluation programs. We speculated that, if a district so intended, it might be possible for them to evolve ways to use findings from testing and evaluation programs to improve instruction.

We defined our inquiry around an initial set of commonsense questions. Were some districts interested in linking testing and evaluation with instruction? If so, what were they doing, and why? What impact did linkage have on classroom practices?

Our definitions for testing and evaluation were not tightly constricted. We defined as within the scope of our inquiry those tests and evaluations which were administered districtwide. Excluded from our definition of testing and evaluation were those tests or evaluations initiated and sustained by teachers for their own purposes or by principals for within-school problem-solving. The large domain represented by the term

instruction included for us all those things which happen in the classroom: --such things as the amount of time spent on subject areas, the materials and tests used by students, the organization of the school day, the teaching strategies used by teachers, the number of adults in the classroom, etc. Likewise, our concept of linkage was inclusive rather than exclusive. Linkage meant relationship --intended as well as actual, unplanned as well as planned, diffuse as well as defined, informal as well as formal.

Selection of sample districts. From RFPs put out by the NIE, from contacts with State Department agencies providing technical assistance to districts, from other research efforts recently completed, we knew that some districts were indeed interested in linking testing or evaluation with instruction. We knew also that there were not very many of these districts. We felt that what each of the districts was doing might differ one from the other.

We started a phone inquiry and collected names of over 40 districts who reputedly linked evaluation with instruction (see Progress Report, 2/80). Categorization of these districts by size, location, ethnicity produced three groups of districts for further investigation. Phone calls to the districts themselves and follow-up visits reduced the number to eight. Further inquiries as to the interst of the districts in participation in our study reduced their final number to four. As a set, we wanted the districts to be representative of other American districts --small, medium and large; urban and suburban; facing common societal problems such as desegregation, influx of minority populations, increases or decreases in population. We wanted a range of socioeconomic conditions, and we wanted districts who themselves said that they were doing something "interesting" in linking testing and evaluation with instruction.

Development of instrument protocols. The need to "bound" our study became quite evident when we began to develop the instruments used to gather information from district respondents. There were differences of opinion among the staff as to the degree of specificity and control that the questions should exercise on the respondent. Some among us argued that we should get the "story" from respondents as they saw it, with only minimum direction from us. Basically these individuals were urging that we go in with a list of topics and, with open-ended questions, obtain as much information as possible; the level of detail and of example would be left to the discretion and judgement of each individual interviewer. There were others on the staff who argued that all respondents should be asked exactly the same questions in the same sequence and in the same order. Advocates of this position emphasized respondent answers were very much a product of the interaction between the interviewer and the respondent. Wording and sequence, therefore, were very important.

Our first set of interview protocols tended in the latter direction. Our curiosity produced over 100 pages of questions. This clearly was impractical. Reducing the questions to a more manageable number meant making trade-offs between breadth and detail. Also, our conversations with district representatives indicated that the length of time available for interviews was likely to be limited to under two hours.

We therefore divided our questions into modules. These modules were to be used differentially with individuals based on their knowledge and longevity in the district. Our original plan for cross-checking respondent perceptions about district functioning had to be abandoned reluctantly.

We pilot tested the modules in a willing neighboring district.

Interviews that we had anticipated would take an hour stretched out to the whole morning, and we still did not gain the depth of understanding that we needed about the complexities of district functioning and dynamics.

We pruned the interview schedules and had them informally reviewed by district administrators and teachers. The final protocols can be found in the Appendices.

FOCUSING PROBLEMS

Selection of respondents. The development of the protocols focused the inquiry into topics such as descriptions of testing, evaluation, instruction, district dynamics, district history, administrator attitudes toward testing and evaluation, teacher attitude toward testing and evaluation, principal and teacher behaviors. Further focusing occurred, often inadvertently, because of the availability of respondents. We intended, in each district, to interview as many as possible of the central office staff and a selection of principals and teachers. This intention was realized, even if imperfectly, in the one small and the two medium-size districts. It was impossible to realize in the large district. In that district, subject to intense pressure because of desegregation requirements, the amount of time and focused attention that officials were willing to give us were limited. We did manage in excess of 20 interviews which, we felt, was sufficient for an in-depth understanding of the complex arrangements relating to testing and evaluation in that district.

With regard to principal and teacher interviews: the selection of individuals to be interviewed was made by our Working Group representatives. Their approach to selecting the respondents varied from

district to district. Although they attempted to ensure for us a representative sample of principals and teachers, they, too, were influenced by the availability and interest of the people they approached.

Use of interview protocols. Although we tried to pre-match individuals with particular protocol modules, we were not always successful. Sometimes respondents talked at length in response to one module, thus preventing us from moving on to the second or third. At other times, respondents were reluctant to speak of sensitive matters and we were unable to record their responses.

It is interesting to note how the interpersonal relationship between interviewer and respondent influenced by the quality of information. Where a few minutes were spent establishing personal credibility and shared interest --especially with teachers-- we were able to get a personal, in contrast to an official, view. Often we had to turn off the tape recorder to continue the conversation. Likewise, the amount of time for the interview influenced the quality of the information. Only late in each interview could we tap into feelings and attitudes. Two-hour interviews are far more satisfactory than half-hour interviews.

Our interviews differed considerably in the extent to which they adhered to the wording and the sequencing of the questions. This had some disadvantages: we have empty spots for some questions for some respondents. We also have answers to questions we did not ask. Nonetheless, this seemed preferable to us than insisting on conformity.*

---

*The four members of the staff who conducted all of the interviews had engaged in extensive discussions about the purposes of the project and of the interview before proceeding. An outline of interviewing procedures is in the Appendices.

AUTHENTICY PROBLEMS

Guba (1978) notes that "problems of authenticity relate to the establishment of bases for trust in the outcomes of naturalistic inquiry." He goes on to ask, "How might the naturalistic inquirer argue to persuade a methodologically sophisticated peer about the authenticity of his information, the categories he derives and the interpretations he draws?"

Guba argues that there are a number of strategies for insuring isomorphism between study data and the phenomena to which they relate. These include: erecting safeguards against potentially invalidating factors such as researcher bias, researcher involvement and researcher variation in recording procedures; establishing "facts" through persistent observation; establishing "facts" through multiple corroboration. As implied in the earlier discussion, these techniques are easier said than done.

Analysis. To each district we assigned a two-person team, and to each interviewer we assigned two districts. This meant that each member of the project staff interviewed respondents in two districts and worked with two other members of the staff. At the end of each day of interviewing, impressions were shared, notes were made, written notes were augmented by listening again to the tapes of the conversations. This was an exhausting and time-consuming process.

In addition to extended discussions among the staff following each set of site visits, we also made successive presentations to our Working Group representatives. This had disadvantages as well as advantages. Among the advantages was clarification and corroboration of our initial impressions. Working Group representatives supplied supplementary details

or alternative explanations. On the other hand, their impressions reflected their point of view. These points of view were argued more persuasively than were the points of view of other district respondents. Since the Working Group representatives were individuals associated with testing and evaluation, they themselves may have had an inflated view of their work and its influence on others, especially teachers.

The analysis of the interview protocols started by identifying everyone who had given us information within the district about a topic. We counted responses where possible and excerpted quotes. We looked for discrepancies and contradictions among respondents. We slowly, through discussion, built up a picture of the district. The original interviewers were responsible for summarizing each of their districts. We then identified themes that were evidenced across districts. Some of these themes are listed in the Appendix.

Nature of observations and findings. We struggled for a long time to find a way of graphically repesenting the categories which would encompass all the details from each district. Some of our thinking during this time is described in the introductory section to the Analytic Framework. We also read what researchers and conceptualizers had to say about school districts as organizations, the dynamics of district relationships, the characteristics of teachers.

We organized the large case study descriptions into loose categories, refined the categories, rewrote the case studies.

The full case studies presented here are one way of representing our data. The shorter analyses are attempts to apply the Analytic Framework to each district. The Observations section is a first attempt at stating the similarities and differences among the districts, drawing upon both the Analytic Framework and conceptualizations in the literature.

PART II.

PERSPECTIVES FROM THE LITERATURE ON SCHOOL DISTRICT USE OF TESTING AND EVALUATION FOR INSTRUCTIONAL IMPROVEMENT

PERSPECTIVES FROM THE LITERATURE ON SCHOOL DISTRICT USE OF TESTING AND EVALUATION FOR INSTRUCTIONAL IMPROVEMENT

In planning and conducting this first year's research we were guided by assumptions and beliefs about the phenomenon we were studying. For example, through previous experience and research work in school districts (Lyon et al, 1978) we knew that evaluation offices had limited impact on local school district operations; we knew that teachers generally did not seem aware of, or make use of districtwide testing results (Yeh, 1979). We knew that school districts, as complex organizations, do not plan for future contingencies very well, and we knew that those plans that they did develop seemed to change during implementation.

As an important part of this first year's work we have sought a clearer focus on our early impressions. Accordingly, we have reviewed literature in the following four areas that directly relate to our interest: school district and teacher utilization of tests and evaluations, school districts as complex organizations, school districts as rational planning entities and teachers as instructional decision-makers.

The selection of these topics for review was consistent with our interest in school districts as the context within which testing and evaluations are administered and utilized.

It quickly became obvious to us that these various literatures were very large and, often, did not focus directly on the subject we were exploring. For example, the general literature on utilization of testing

and evaluation is growing very rapidly --there is an exploding interest in the topic-- yet there is little about the management of such activities within the district context. Also, there are surprisingly few studies on the planned change process at the school district level. Much of our review of that topic is built around general theories of complex organizations and the rationality of their planning processes where there is much conceptual work, but limited empirical data.

What follows, then, is a selective review of literature to help us sharpen our focus and thereby gain a better grasp of the phenomenon we are examining.

As one might expect, there is considerable conflict in the literature about some topics - the answers are not always clear. In this review, we will not try to resolve those conflicts, but we will try to interpret the conflicts in terms of our interest. For example, there is disagreement over how much autonomy teachers really have in the classroom. We will consider that conflict in terms of the extent to which teachers have autonomy to make decisions regarding their use of tests and evaluations in classroom instructional decisionmaking.

This review is constructed around the following four areas:

1. A description of the extent to which school districts, principals and teachers currently use tests and evaluations in instructional decisionmaking.
2. A description of selected characteristics of school districts as complex organizations - and the implications these characteristics have for district use of testing and evaluation in instructional decisionmaking.
3. A description of the role planning plays in school district change.
4. A description of teacher characteristics and role behavior and the implications these have for district

use of testing and evaluation in instructional decision-making.

1. School district and teacher utilization of tests and evaluations

Within the past five years, there have been a small number of empirical studies investigating the extent to which evaluations and tests are useful to the people who are directly responsible for the education of children. These studies have been spawned by the growing realization among those who commission testing programs or evaluations that the benefits they hoped would flow from these forms of data collection and analysis have not materialized. Some investigators believe that these hopes were --and still are-- ill conceived. Others want to look carefully at the many facets of the knowledge utilization problem with some expectation that testing and evaluation can be fine-tuned into agents of instructional improvement.

We are here interested in empirical studies which identify factors facilitating or impeding district or teacher use of evaluations and tests.

It is clear that the important differences in purposes, characteristics and implementation among testing and evaluation programs affect the way the results are regarded. Alkin, et al (1974) found that a formative or improvement orientation by the evaluator to his/her role made the project director's acceptance of evaluation findings more likely. Alkin, Daillak and White (1979) identified a more detailed set of contextual, professional and interpersonal factors which appeared to affect the evaluator-project director relationship. Daillak (1980), by shadowing central office evaluators as they worked with school

personnel, concluded that the provision or withholding of information during data collection as well as after data analysis, is a political more than an interpersonal phenomenon. Individual or group interests at different levels within the district and between the district and outside agencies often deflect evaluations from their manifest function of identifying program effects.

The use to which Title I evaluations were put by local educators was investigated in 15 districts in 6 states by David (1978). Her conclusion was that district officials and teachers regard evaluations primarily as meeting federal reporting requirements. These evaluations were therefore associated with accountability rather than seen as potential sources of useful information. The accountability threat, added to the narrow way in which the evaluations were defined - i.e., as dependent on standardized test scores - contributed to a minimal level of local use even in districts purportedly interested in evaluation.

Board members and Superintendents surveyed by Tomblin and Harding (1980) in 296 small school districts reported a high degree of evaluation activity in three areas: monitoring of the ongoing programs; assessment of student progress; accountability to outside agencies. However, two-thirds of the Boards report that they do not have a policy for program evaluation; the remaining third said that it was either a state-mandated or a district policy - but were vague as to the details of that policy.

The terms testing and evaluation are used interchangeably by many district officials and teachers. What they refer to by these terms is non-teacher-initiated tests mandated by an agency outside the classroom. Many survey and interview studies have found that information from these

activities is not used by teachers (Goslin, 1967; Stetz & Beck, 1979; Boyd, et al, 1975) because such information was seen as either irrelevant, redundant or inaccurate. A CSE survey of teacher use of tests (Yeh, 1980) showed that teachers prefer their own tests or tests embedded in the curriculum materials they use. "The more closely a test is related to instruction, the more likely it will be used." (p 36)

These findings imply that instruments developed without regard to the specifics of teacher needs or curriculum applicability --whether they are called testing programs or evaluation efforts-- are likely to be ignored if possible and barely tolerated if necessary. It would seem that strong incentives would be needed to encourage district personnel or teachers to change their behaviors based on the evidence produced by these instruments.

2. A description of selected characteristics of school districts as complex organizations and the implications these characteristics have for district use of testing and evaluation in instructional decisionmaking

We have selected four characteristics of school districts as complex organizations that in our opinion likely influence the way and the degree to which school districts use testing and evaluation in instructional decisionmaking. The four characteristics are:

a. Domestication and institutionalization

b. Goal diffuseness

c. Permeable boundaries

d. Loose coordination

a. Domestication and institutionalization. By domestication of school districts, (Carlson, 1964; 1965) we mean that they are subject

organizations with little autonomy in certain critical matters. They have very few options about choosing their clients (students) and the clients in turn are obligated to attend the schools run by the district. (Some students, of course, can select private schools, but the percentage who can do so has generally been quite small in most communities.) What is more, the schools' financial support is externally determined, minimally guaranteed, and does not vary with the schools' performance. (Carlson, Gallaher, 1965)*

One consequence of this domestication is that the school or its parent organization --the district-- need not "prove its worth" on any regular basis. That is, unlike many private sector organizations, the continued flow of clients and funds is relatively independent of the quality of services. Consequently, the school district's technical services - i.e., instruction - receive less sustained attention than they would in other organizations, public or private, where the relationship of the technical core to organizational survival is more clearly evident. Viewed from this perspective, there does not appear to be a strong incentive for school districts to utilize testing and evaluation in improving the instructional technical core (Meyer, 1977; Niedermeyer, 1979).

Indeed some (Meyer and Rowan, 1978) argue that it is detrimental for school districts to link testing and evaluation to instructional

---

* Recent developments in some school districts, e.g., desegregation efforts, and in some states, e.g., diminishing state tax support, may modify this domestication characteristic. More and more parents are finding ways to send their children to private schools or move into neighboring school districts. In spite of these trends, however, it is likely that most school districts can still be accurately described as domesticated.

decisionmaking. They reason that schools, like public libraries and national parks, have been institutionalized. They are valued intrinsically, not because they have survived competitively or because they have been able empirically to demonstrate their social contributions, but because the public believes prima facie that schools have high social worth. This being the case, school districts have no need to demonstrate their technical worth in terms of student performance; furthermore, it may be that close public scrutiny of student test results will reveal that the schools have, in fact, a weak technical core and that they are not doing a very good job of teaching. This may already have begun to cause a serious breach in public confidence, with a resulting decline in public support, both in terms of dollars and pupils.

Some might counter this institutionalization argument by asserting that tests should not be used for judging the organization's effectiveness because they are inadequate and narrow measures of pupil performance. Meyer and Rowan (1978) however, observe that tests are devices routinely used to make very important decisions regarding students' career options, e.g., college aptitude, academic or vocational tracking. Why are tests valid enough to use in judging students, but not valid for judging teachers' or the schools' performance? Given their institutionalized characteristics and the risk that may accompany revealing the schools' weak technical core, it would appear unlikely that school districts will readily and easily adopt or develop a program to link testing and evaluation with instructional decisionmaking - indeed it may be in their best interests not to do so.

b. Goal diffuseness. Assuming the domestication and institutionalized characteristics of schools described above, one can expect that there will be considerable lack of precision in defining the school district's goals or the goals of individual schools. While some have noted that there seems to be general agreement on what society expects of schools, (Goodlad, 1979), there is considerable disagreement in determining which are the most important goals and whether or not the schools have satisfactorily achieved those goals. In support of this position, Boyd (1979) has pointed out the many additional activities and responsibilities assigned to schools over the years --for nutrition, medical services, counseling, etc. Accompanying these additional tasks have been interest groups, each pressuring the schools to give higher priority to selected types of students, e.g., economically deprived, handicapped, bilingual, gifted. Schools seem to have neither the authority nor the means to logically order their priorities and set about a carefully planned course of action to achieve their goals (Elboin Dror, 1970). This results in goal diffuseness -- the district's and the schools' central core activities are neither agreed upon nor consistently pursued.

Given unclear goals, schools tend to point to inputs and ceremonies as measures of school effectiveness (Meyer and Rowan, 1978; Weisbord, 1978). For example, the instructional problems resulting from a high concentration of low-income pupils in a school are resolved by providing additional funds and other resources in the hope that they will "make a difference." Surprisingly little attention is given to analyzing the way alternative instruction might remedy the pupils' academic deficiencies.

What is more, public attention is directed towards ceremonial and symbolic factors, e.g., use of approved texts, supervision by credentialled professionals, even though there is little evidence that "more of the same" will make much of a difference in student learning.

One implication of this goal diffusion and ambiguity is that, lacking clear priorities and given continued pressure for shifting established priorities, it is unlikely that districts will develop a coherent plan to link testing and evaluation with instructional improvement. They know that it is difficult to hit a moving target.

c. Permeable boundaries. From an open system viewpoint (Katz and Kahn, 1978; Rogers, V., and Rogers Argarwala, R., 1976), school districts are engaged in a continuing interaction with their relevant environments, both proximate and distal, e.g., community, state and federal agencies. Organizations vary with regard to the intensity and frequency by which the external environment can breach the organizational boundaries and influence the organizations' direction and internal operation. School districts, as organizations, have very permeable boundaries.

As has been noted above, school districts usually protect their instructional technical core from close public scrutiny. However, the public has long had influence over other vital organizational matters, especially capital outlays and student regulation, e.g., use of corporal punishment. Lieberman (1977) has observed that, at the local level, schools are among America's most publicly influenced institutions. While influences from the state and federal levels have been less salient, there is an accelerating trend towards their regulation of school district activities. This has come about in part because school

districts have become dependent on federal and state financing to aid them in addressing what are considered pressing social problems, e.g., reducing racial isolation, mainstreaming handicapped pupils, etc. Regulation and some control almost always accompany such funds. This results in ever-increasing breaching of the school districts' boundaries and a resultant lessening of local control. (Halperin, 1980; Kirst, 1980)

Districts differ, of course, with regard to how they are affected by this external regulation. Some small districts have very few students who are the targets of federal or state attention. Consequently, these districts remain relatively free of external regulation. Other districts, larger or in urban settings, have many such pupils and consequently receive considerable funding - and regulation. These differences account for diffuseness in the level and intensity of district testing programs. For example, Title I regulations require normal curve equivalent tests as the means of documenting program effectiveness. Courts often specify the type of documentation and data they will accept as evidence of district competency in providing equality of educational opportunity. Given the power of the legislative, executive and judicial branches of the federal and state governments to impose stringent sanctions for lack of compliance (e.g., withdrawal of funds, court-ordered busing) some school districts are having to publicly discuss and justify the effectiveness of their technical activities --that is, their instructional program. The breaching of boundaries in this area is a relatively new phenomenon. The effect of this trend is not viewed by observers as necessarily ameliorative or logical (Wise, 1977). It does seem possible, however, that such penetration, if it continues, may have far-reaching consequences

on other characteristics of schools, especially those of goal diffuseness and loose coordination.

d. Loose coordination. School districts are typically analyzed in terms of the standard bureaucratic model (Bidwell, 1965; Corwin, 1974; Weber, 1974). They share common bureaucratic properties with other complex organizations, such as hierarchy of authority, defined rules and regulations, job specialization. But, they also have distinguishing features because they are public educational agencies.

A current and widely held view of public school districts is that, as compared to some other bureaucracies, they are "loosely coupled." (Meyer and Rowan, 1978; Weick, 1976.) This means that the linkages among individuals operating at the various levels of the system --central administrators, principals, classroom teachers-- are weak. As a result, directives, orders or regulations from the central administration can be ignored or incompletely implemented at lower levels. The intent and meaning of messages often are diffused or altered as they pass up or down through the various administrative levels.

While loose coupling is a useful metaphor for describing the authority chain in public school districts, the work of several researchers (Abramowitz, 1978; Daft and Becker, 1978; Downs and Mohe, 1976; Lortie, 1977; Spence et al, 1978) suggests that this loose coupling analysis should be applied differentially and selectively to specific aspects of school district functioning. That is, some decisions and regulations seem to be tightly coupled --enforced very directly and effectively. For example, central administrative policies regarding pupil control (e.g., use of corporal punishment) or school

scheduling (e.g., beginning and ending times for the school day) are closely monitored and fully implemented in all schools and classrooms. However, in other spheres, e.g., teaching strategies or classroom management, there appears to be loose coupling (Firestone, Herriott, Meyer, et al, 1978; Meyer and Rowan, 1978). Central administrators here rely on "confidence and good faith" with regard to teachers' instructional decisionmaking. The tradition of teacher autonomy (Lortie, 1975) reinforced by increasing teacher union power, weakens the ability of central administrators to influence routine teacher classroom behavior and instructional decisionmaking. This lack of influence has not resulted in organizational anarchy (Cohen, et al, 1972). Rather, the relative independence of classroom teachers and their loosely coordinated instructional activities may only mean that the instructional actions or decisions of one teacher have little impact or importance for other teachers. The consequences of this lack of coordination on children may be salutary --exposing them to diversity-- or harmful. There is little empirical evidence on this point.

Schools and classrooms, in terms of instruction, can be thought of as being federations or collections of zones of influence. (Lortie, 1977; Miles, 1980.) Decisions about technical matters --i.e., instruction-- rather than being centrally determined, widely disseminated and carefully enforced are, instead, parceled out to those who have the immediate and direct responsibility for action.

3. Planned change in some districts

We have pointed out that school districts as organizations differ from other complex organizations in some ways, and share in some common

characteristics. One important characteristic of complex organizations is that they constantly change (March, 1980). Observers of organizations often overlook this "change" characteristic and, instead, they stress the rigidity of organizations. But, as March reminds us: "Organizations are like the weather, changeable enough to animate casual conversation, steady enough to make the conversations repetitious." (p 5)

This combination of change and constancy is observable in school districts. School districts as an organizational type, however, have the other attributes described earlier that influence their functioning, e.g., domestication and institutionalization, goal diffuseness, permeable boundaries, loose coordination. Thus, school districts may be described as complex organizations "that are open to their environment and have the capacity to change while maintaining their basic system properties." (Berman and McLaughlen, 1979, p. 7.)

Over the years, educational researchers have been interested in developing or investigating methods of systematically planning and implementing change. Much of that research literature has focused on small-scale changes, e.g., changes in classrooms, or individual schools, or developing and implementing projects. Relatively little has been directed toward studying planned change at the school district level (Pincus and Williams, 1979). There is a weak empirical basis upon which to make confident observations about the change process at the school district level.

Presently the dominant model for viewing planned organizational change is the technological viewpoint (Clark, 1980; House, 1979). This

view holds that there is a systematic logical process which, if followed, will maximize the organization's chances of changing and reaching desired goals. The steps in this logical process usually are framed as: determine goals, develop plans, review programs, and evaluate results. (Clark, 1980)

If the goal is not realized, or if the plans are not implemented as hoped, then rational planning advocates often assume some technical adjustment can be made so that this logical planning process will work. Seldom, it appears, do rational planners consider that the rather widespread failure of this planning model might be because this approach has basically faulty supporting assumptions. (Clark, 1980)

Some (Clark, 1980; Berman and McLaughlin, 1979, Lindbloom, 1958; March, 1980) have sharply questioned the applicability of this logical planning process to accomplish significant change in organizations such as school districts. They argue that characteristics such as multiple subsystems, powerful informal organizations, goal conflict, uncontrolled external environmental influences, make such logical approaches intrinsically unworkable. They argue that it is unlikely that any rational planning model can be sufficiently powerful or comprehensive so as to accommodate these various influences and thereby result in logical, carefully sequenced program implementation.

House (1979) argues a better understanding of the change process in complex organizations can be realized if two other frameworks --the political and the cultural, are employed. The political framework emphasizes influence, power, persuasion, coercion, and negotiation in assesing organizational behavior. It assumes that complex organizations are made

up of many groups with often conflicting views and varying amounts of power and influence, and that planned changes are transformed as they run the political gauntlet. It follows that plans will quite naturally and expectedly be changed as they move from conceptualization to implementation.

The cultural perspective views the many parties involved in an intended change effort as members of various cultures, each with its own set of norms and expectations. Each time intentions move from culture to culture, they are reinterpreted and transformed not so much in power and political terms, but in value terms. Advocates of this approach see the schools in evolutionary or ecological terms (Farrar, De Sanctis and Cohen, 1979; Goodlad, 1975). Intentions will be transposed during implementation in ways that reflect the values of the individuals who belong to various cultural groups. Consistent with this evolutionary approach is the observation by Pincus and Williams (1979) that the pressures on and processes in innovative school districts seem to keep them on an irregular cycle of advance and retreat.

As House (1979) persuasively argues, the change process in complex organizations can be most productively studied when all three frameworks (technological, political, cultural) are used. Districts can and do benefit from planning future activities in a logical and sequenced way. However, school district planning can be made more useful and better understood if political and cultural perspectives are kept in mind.

4. Teacher characteristics and role behavior

One of the desired outcomes of an effective T/E/I linking subsystem is that teachers would increasingly use the findings of districtwide

testing programs and evaluation studies as a basis for thinking about students' learning and their own classroom actions. Here, using what House (1979) would consider a cultural perspective, we review relevant research on teachers and the teaching role to assess the likelihood that they would utilize test results and evaluation in instructional decisionmaking and subsequently in determining their teaching behavior.

A number of observers and researchers have noted the comparatively weak preparation teachers receive in either the technology of teaching or the technology of testing. Teachers are often not well trained either in remedying particular student deficiencies or in the interpretation of test scores. Because teachers rely on their own experience in learning what they regard as the craft of teaching, they tend to develop their own individualistic teaching and evaluation methods in which they express considerable confidence. In Lortie's terms, "the total induction system is not highly developed" (1975).

Teachers accrue a considerable amount of data about students from their own testing activities and from prolonged exposure to students and extended observations. To many teachers the most critical evidence of their own effectiveness is their own perception of student behaviors. They see noncompliant behavior and lack of enthusiasm for study as indications of poor learning, while enthusiasm for the teacher and subject matter is seen as indications that students are learning and enjoying their work (Lortie, 1975). Sarason (1971) notes, however, that teachers experience considerable tension between satisfying individual needs and "covering the material."

Miles (1980) reports that 50% of the teachers responding in the NEA's 1979 teacher opinion poll said that "most of the really important

aspects of student progress are not measured by tests" (p 98). Moreover, Lortie (1975) suggests that teachers discuss achievement test performance "as if they are uncertain of the tangibility of measured gains or the rightfulness of their claiming credit for them" (p 128).

Even if teachers are persuaded by standardized test results, there is often a gap between identifying the pupil's shortcomings on the one hand and determining appropriate instructional activities on the other (Lortie, 1975). In other words, there are few accepted cause and effect relationships within the profession that guide teachers from assessment to appropriate action.

Teachers have considerable autonomy in deciding whether or not to use standardized tests and evaluation in instructional decision-making within their own classrooms. This autonomy is the result of the limited amount of classroom supervision provided by school principals (Abramowitz, et al, 1978; Packard, et al, 1978). Because of the "egg carton" design of most school buildings, teachers generally perform isolated from their colleagues and beyond the day-to-day supervision of principals (Goodlad and Klein, 1970; Lortie, 1975; Meyer, et al, 1978). Also, there are few incentives or other means available to supervisors who want to increase the use teachers make of test scores and evaluation findings. Experienced teachers have tenure and cannot be dismissed if they do not comply with the district program and utilize tests and evaluation results. At the same time, the districts have neither financial resources nor other rewards to offer as inducements.

Given these realities, districts hoping to induce teachers to depart from their usual methods of operation must do so with the claim

that testing and evaluation will help teachers solve pressing day-to-day problems, e.g., how best to group pupils for instruction, how to diagnose learning difficulties, how to enhance pupil learning so that the teacher can "cover the material" and win psychic rewards from the pupils and the supervisors.

In summary, teachers by training and experience and incentives would not seem to readily turn to tests and evaluations that they themselves did not authorize as data for making instructional decisions or planning classroom activities (Deal, 1975). Districts have few coercive methods or extra compensations they can use to increase teacher utilization of test and evaluation. Likely, any administration that hopes to induce teachers to use testing and evaluation must appeal to their programmatic classroom concerns.

PRELIMINARY ANALYTIC FRAMEWORK OF T/E/I LINKING SUBSYSTEMS

PRELIMINARY ANALYTIC FRAMEWORK OF T/E/I LINKING SUBSYSTEMS

Development of the Framework

After collecting interview data, we wrote a first draft case study for each district which provided us with a summary of our respondents' answers to our questions. We next attempted to identify themes or categories of relevance across districts from our data and from the literature review. These themes were to serve as the basis for analyzing what the data meant. It seemed to us that although the details of what each district was doing differed, we should be able to identify functional equivalents across districts. For example, one district had formal committee-based ways of communicating; another operated on the basis of shared and implicit understandings. In each case, however, the function of communication was carried out in a sufficient and satisfactory manner.

We then looked for an analytical framework that would allow us to recategorize data by topics and to analyze them according to specific concepts. We struggled to define the relationships between and among the categories and topics. We searched for extant analytical categories or frameworks that might describe what we had found. We discovered none that seemed to fit. For example, we considered the standard planning or development model with its sequential processes --define the problem, search for solutions, select the best solution, implement the solution, evaluate the results, and then recycle the process. This simply didn't apply, even if we defined as the "problem" the linkage between testing, evaluation and instruction. We noted that many researchers of credence argue that such a rational, sequenced process seldom, if ever, is actually utilized in planning and implementing sound programs (Lindbloom, 1958, Clark, 1980).

We considered a related construct that applied to the stages or phases of educational change, i.e., initiation, implementation, evaluation. It, too, did not apply even if we defined the linking we were looking at as an "innovation." Generally, in our districts linking tests or evaluation with instruction was not recognized as a change or innovation until after it was almost fully developed, if at all.

We considered Lewin's Force Field Analysis but our respondents' use of testing or evaluation in relation to instruction could not be conceptualized as a clearly defined or planned "program" that was buffeted by forces acting either for or against its use.

The most useful approach for us seemed to be open systems analysis. However, the complexity of urban school districts is such that a thorough understanding of these organizations so as to locate our linking subsystem within them in a full open systems sense, was beyond the scope of our work.

We, therefore, inductively developed our own analytical framework which, admittedly, utilizes some categories or concepts of other models (for example, that of "relevant environments" from the open-systems approach).

We have deliberately referred to these preliminary categories and analytic concepts as a framework rather than a model, paradigm, or theory. We are as yet unsure of our analytic tools and of the relationships or dynamics within the framework so model or paradigm seems an inappropriate term; the framework has no predictive function - thus, it cannot be considered a theory. It is, simply, a way of arraying the data in order to gain insights into the dynamics of the linkage subsystem. Refinement and certainty will both increase during the next two years.

Description of the framework

The proposed framework describes and provides the basis for analyzing a subsystem within a school district that connects testing and evaluation activities with instructional improvement. It is, therefore, an attempt to describe what we mean by "linkage." The T/E/I linking subsystem that we propose is a construct --that is, it is not an organizational entity such as a division or a unit that is clearly identified within a district but it does have properties which can be described and analyzed across districts. It is not necessary to an understanding of this framework to define precisely the terms evaluation, testing or instruction. Common usage suffices for the present - except to say that we exclude from our definition those testing and evaluation activities which are teacher-initiated within the classroom. We refer only to a set of testing and evaluation activities which are initiated by someone other than the teacher.

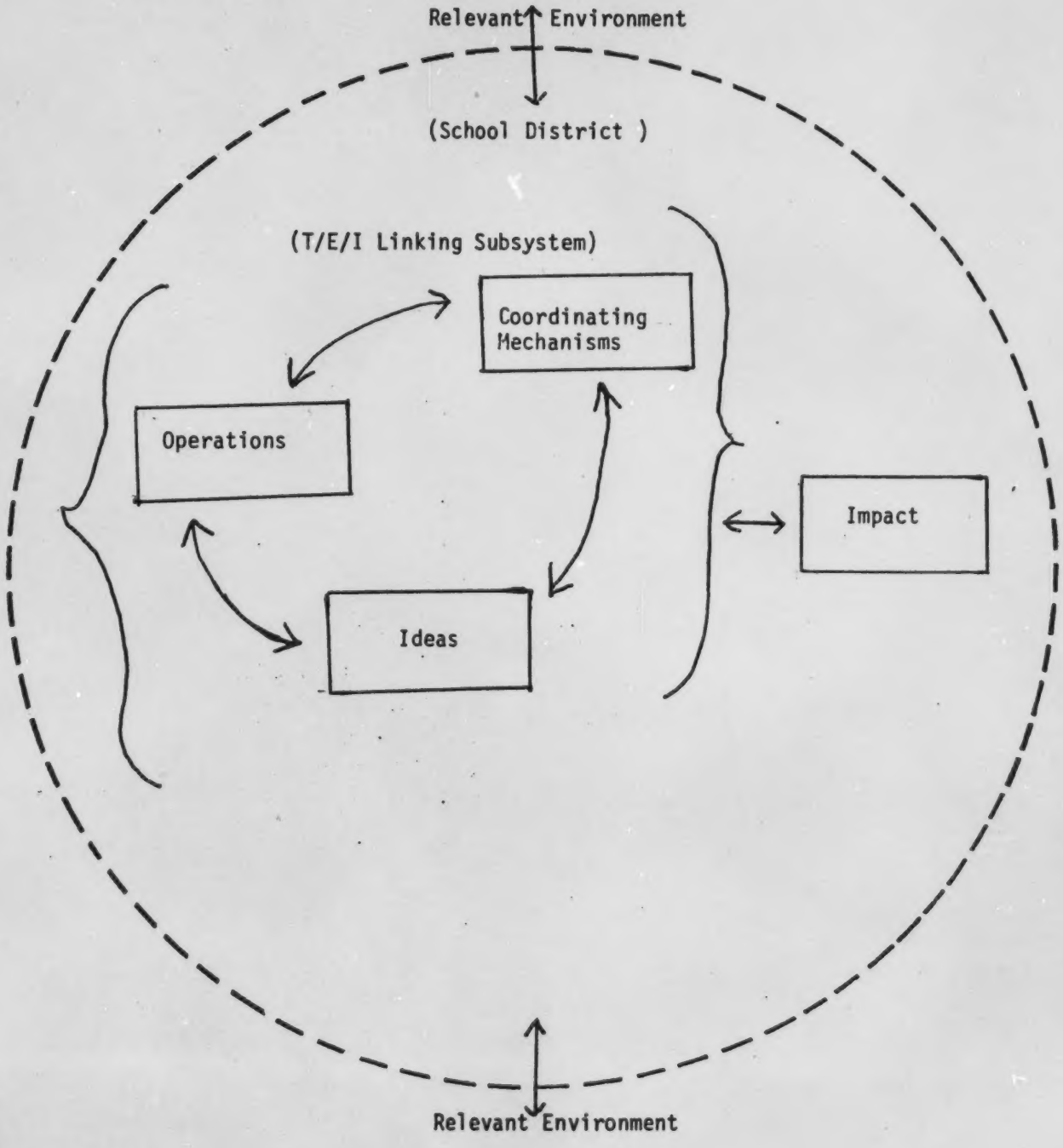
The framework consists of a set of categories, each with subtopics. As noted earlier, the choice of categories and subtopics was dictated by cross-site analysis of our four districts. In addition, within each category, we include dimensions on which to characterize the categories and subtopics. Finally, the framework suggests relationships among and between the categories.

The categories, subtopics and dimensions are first outlined, then diagramed, and finally discussed on the following pages.

Outline

(Categories & Subtopics)	(Dimensions)
RELEVANT ENVIRONMENTS FOR T/E/I	
Societal Community Professional Intra-District	Can be characterized by degree of saliency and penetration.
IDEAS ABOUT T/E/I	
Evolution	Can be characterized as pro- to reactive, salient to adjunctive, planned to unplanned.
Content direction management	Can be characterized as potent to weak, comprehensive to fragmentary.
OPERATIONS OF T/E/I	
Number and kind of components in subsystem	Can be characterized in terms of degree of involvement.
Quality of staff performance	Can be characterized in terms of level of competency.
COORDINATING MECHANISMS FOR T/E/I	
Formal Informal	Can be characterized by degree to which functions such as communication, commitment-building, decision-making are fulfilled.
IMPACT OF T/E/I ON CLASSROOM ACTIVITIES	
On teacher knowledge, attitudes, skills	Can be characterized by type and level.
On teacher-initiated activities On nonteacher-initiated activities	Can be characterized according to topic; instructional, interpersonal.

T/E/I Linking Subsystem
Preliminary Analytic Framework



Discussion

1. Relevant environments.- Taken from general systems theory, this category refers to those conditions or events external to the district and/or to the T/E/I linking subsystem that have influence on individuals within the subsystem. These environments can be defined in many ways and can be as inclusive as seems appropriate for analytical purposes. Examples of relevant environmental influences are: size and wealth of the district, district socioeconomic and racial characteristics, societal or political events such as desegregation orders, population transiency, federal or state mandates and requirements. These relevant environments form into configurations that are unique to each district and, in turn, influence both the district and the T/E/I subsystem as to what can or must be done. They both provide the context for organizational action and are in turn influenced by organizational action.

2. Ideas.- This category refers to ideas about testing, evaluation and instruction, their relation to one another and to the broader concerns of the district held by individuals in the district. We have chosen the broad, encompassing term "ideas" instead of more commonly used and narrower words such as "policies," "objectives," or "plans," because these latter nouns do not include the images, wishes, hopes that are not yet formalized. We describe ideas in two ways: in terms of their evolution and in terms of their content. Their evolution may be characterized in terms of their sequence - (e.g., planned to unplanned, phased to haphazard) and in terms of their intentionality (e.g., primary to adjunctive, pro-active to re-active). The content of the ideas can be further divided into ideas about direction and ideas about management.

Ideas about direction have to do with such topics as: where testing and evaluation fit into district problems and activities, which tests or evaluations to focus upon and why, the relationship between district's testing or evaluation activities and the district's instructional continuum, the relationship between in-service training and testing and evaluation, etc. Ideas about management have to do with: ways in which components other than tests and evaluation, e.g., in-service training, should support these components and instruction; ways in which the components can be coordinated; how the entire effort can best be connected to classrooms.

Ideas can differ with regard to at least two characteristics: potency and comprehensiveness. An idea's potency is high if many people share the idea, think it is important, and agree with it; its comprehensiveness is high if the factors or components are coordinated, consistent and complete.

3. Operations.- By operations we mean those materials, activities, and resources or organizational units that make up the T/E/I linking subsystem. These would include: specific tests administered (e.g., CTBS, State Assessment Tests, District Proficiency Examn, CRTs); specific evaluations performed (e.g., Title I, Title VII, School Improvement); staff development (e.g., teaching techniques, subject-matter content, technical knowledge of tests); curricula (e.g., district or school continua); materials; supervision; resources (e.g., evaluators or testing specialists, learning specialists, department chairpersons), financial resources; personnel.

These operations can be characterized in terms of their degree of involvement in the T/E/I subsystem and their level of competency. That is,

what operations are present, and are they connected in a way that supports the T/E/I linking subsystem; and is the performance of each operation of high quality.

4. Coordinating mechanisms.- By coordinating mechanisms we mean those formal and informal activities and structures the district uses to maximize staff commitment to, and staff communication about the district's T/E/I linking subsystem and the ways in which decisions are made about it and within it. These can include both formal and informal organizational mechanisms. Formal mechanisms might be cross-functional committees, special purpose committees, periodic meetings of constituent groups, special bulletins, school-site councils.

Informal mechanisms typically flow from, and can be influenced by, the formal organizational structure. These include the wide variety of cliques and other small friendship or occupational groups that form in all organizations. These groups can exert either positive or negative influence and any effective formal organization must account for their formation and potential influence. Also, the district can develop an ethos of efficacy or high morale which can affect staff beliefs and performance.

5. Impact. By impact, we mean the extent to which the T/E/I linkage subsystem is having the desired or expected effect in classrooms. This effect or impact can be considered in two ways: the influence that tests or evaluations have on teachers' thinking and intentions, i.e., are they aware of the tests or evaluations and of the data from them; do they value the tests and evaluations, do they try to use the data for any purpose; and the impact tests or evaluations have on classroom activities, whether these activities are teacher-initiated, e.g.,

grouping, diagnostic use of tests; or whether these activities are externally-initiated, e.g., district-mandated use of a particular teaching method such as Distar, district-organized remedial programs, etc.

Relationships.- The two-way arrows in the diagram indicate the dynamics of the framework, i.e, that influences flow in many directions. The circular, nonlinear or nonhierarchical arrangement of the categories (Ideas, Coordinating Mechanisms, and Operations) indicates that districts may start with or emphasize one or more categories before others and that categories do not follow any particular sequence, e.g., ideas do not necessarily precede the other two components.

The relevant environments interact with the district and its subsystems. Changes in the relevant environment, e.g., considerable and rapid change in populations, community socioeconomic and racial makeup can have a profound influence on the organization and on the subsystem. Similarly, changes or events in the organization or subsystem, e.g., declining test scores, can influence the environment in ways which in turn, influence the organization.

The delineation of the relationships among and between categories are, at this point in our work, quite primitive and speculative. We have observed how these interactions occur in some instances. Much of our work in the next year involves clarifying these relationships as well as refining the other elements of the framework.

CASE STUDIES OF FOUR DISTRICTS*

*Please note that the names of the school districts and respondents have been changed so as to maintain confidentiality. Some demographic details have been altered for the same reason.

BAYVIEW UNIFIED SCHOOL DISTRICT
1980

BACKGROUND INFORMATION

The Bayview Unified School District is a medium-sized district in a community with a population of about 100,000. It is a growing community with both the numbers of minority and low income students increasing regularly. About 7,000 new housing permits have been recently issued representing most of the unused land in the city limits.

Bayview's student population numbers 14,000, 52.7% of whom are minority. Of those, black students represent the largest single minority, approximately 30%. There is a large Filipino student population, approximately 11%, and a wide variety of other minority populations in smaller numbers. The socioeconomic status of Bayview's student population is extremely diverse. For example, recent data indicate that 7% of Bayview's third-grade students come from professional families, 17% from semiprofessional families, 48% from skilled/semiskilled families, and 30% from unskilled or welfare families. The percent of students from families receiving AFDC (Aid to Families with Dependent Children) is high throughout the District and recently ranked at the 88th percentile in comparison with other of the state's school districts.

There are 22 schools in the Bayview School District: 16 are kindergarten through 6th grade, 4 are 7-9 grade junior high schools, and 2 are senior high schools. In addition, the District has 8 children's centers and preschools, a school for trainable mentally retarded children, a continuation high school, and three alternative programs for youngsters in grades 7-12, and an adult school. Also, the Community Education Department serves post-high-school-age citizens.

The District's 1979-80 budget is $28,790,790 and the state has classified the District as a "low wealth" school district. In spite of this, Bayview has the reputation of being an innovative district. This stems from its efforts for the past six years in organizing staff development programs for elementary and secondary principals and teachers. The District currently operates a state Professional Development Center Program, a federal Teacher Center Program, a federal Teacher Corps Program, a Right-to-Read Program, and an Emergency School Aid Act Program. The Professional Development Center Program is coordinated with the Title I programs in providing staff development services.

RECENT HISTORY OF BAYVIEW

Bayview, a community of 100,000 located about 50 miles north of a major metropolitan center, is still a small town in spite of the recent influx of both minority population and former city dwellers. A great number of adults can remember attending Bayview schools as children. The next few years may see it become primarily a bedroom community with suburban commuter characteristics; or a small city with the problems accompanying rapid growth.

The Bayview School District, with 16 elementary schools, 4 junior high schools, 3 senior high schools, approximately 700 teachers and 14,000 students has, over the years, mirrored in its purposes and its functioning the general trends arising from within its community, within the country and within the profession of education.

During the immediate postwar period and well into the 1950s, Bayview experienced an enormous growth in population, with workers from the East coming to seek their fortunes in its shipyards, fishing and other industries. The presence of the military expanded as Bayview's naval base enlarged. The housing industry flourished. Expensive housing replaced cheaper wartime construction. Twelve new housing projects went up on the east side of town within a few years. Bayview acquired a slightly unsavory reputation with the neighboring communities - as a wide-open brawling heterogeneous town where quick money could be made. The small

central office of the District coped with overcrowding as best they could - building new schools, giving a double session, expanding the pupil-teacher ratio. Attendance in some of the classrooms was as high as 45 children. The governance of the District was centralized with the central office subject area coordinators overseeing the subject matter content in each of the schools.

During the decade of the '60s, there were substantial changes in the way the District conducted its business. These changes were partly due to overall student political and social unrest which manifested itself in Bayview -- as in other American towns -- in racial incidents, formation of gangs, political activity related to school desegregation (most of Bayview's black students went to Bayview High School -- but the white students held the leadership positions in the student government.) Additionally, the changes in the District were stimulated by President Johnson's Great Society programs, especially the 1965 Elementary and Secondary Education Act which for the first time in American history committed the federal government to massive amounts of spending for compensatory education and school improvement purposes. Finally, the sixties were a time of intellectual ferment about the processes of schooling. In the late fifties attention had been paid to curriculum revision, especially in math and science, in hasty response to apparent Soviet superiority in Sputnik-type space technology. By the late

sixties, the search for educational panaceas had turned to classroom management techniques such as individualized instruction, team-teaching, multiage grouping and learning centers. Many new teachers exposed to these ideas at the nearby university, came to Bayview.

The Superintendent who was brought to Bayview in 1970, Bill Grady, was the first person in a long time who was imported from the outside - a non Bayviewer. He has been described as a "tightfisted, very directive, very innovative person." His two-year stay in Bayview introduced two new ideas: decentralization of curricular and instructional decisionmaking to the schools and encouragement of central office staff to write proposals for federal money.

In 1972, Fred Miller, a new Superintendent, also from outside the District, although with previous experience in the state, was appointed. His abilities and interests lay in the area of school desegregation and intergroup relations. He also encouraged schools to develop mechanisms for school-site coordination and staff development -- a team-building approach to school-site governance. Furthermore, he continued the practice of seeking outside support for activities the District wanted to initiate or maintain.

The Superintendent's attitude provided the fertile atmosphere within which the faculty in one of the District's predominantly middle-class elementary schools -- Hilltop -- could flourish. Mike Mitchell had been in the Bayview schools for many

years as teacher and principal, As part of his search for ways to improve instruction at his school, he participated in workshops run by Madeline Hunter of UCLA. For him, these workshops provided a framework for diagnostic prescriptive teaching leading to individualized instruction. What was more, Madeline Hunter inspired him to believe that he could make changes at his school. Mike went back to Hilltop resolved to move towards such a system. He found two teachers highly receptive to the Hunter ideas. Together, they attempted -- without a whole lot of planning or preparation but with enthusiasm to team-teach in their 6th grade classes. A third person--a special education teacher--worked with them. His interest was to give his EMR children a regular classroom experience whenever appropriate. The teachers worked hard. Experience led them to believe that they themselves, and other teachers as well, needed more training in the critical elements of teaching (e.g., diagnosis and prescription, use of valid principles of learning, etc.). They worked together in in-service courses, weekend workshops and in teachers' meetings to increase their own teaching effectiveness.

Hilltop's work on teaching effectiveness was so successful in the opinion of the Superintendent that it should be fostered district-wide. Consequently, Mitchell and one outstanding teacher - Jane Daniels - were moved into the central office. With the departure of the Superintendent in 1979, Mike and

Jane's influence has increased. Mike is the Director of Instruction. Jane is the Coordinator of Staff Development. Other teachers for Hilltop School have moved into the central office or become principals.

They have acquired resources to support their perspective that staff development in techniques of teaching is the way to improve instruction. Over the past five years, the District has received $300,000 in outside funding for the Professional Development Center and the Teacher Center.

Staff development is the core activity which stimulates other ideas within the District and around which other functions coordinate. The extent to which testing and evaluation will become salient factors in improving instruction would seem to depend, in large measure, on the attitudes of the leadership group.

WHAT BAYVIEW IS DOING TO LINK TESTING AND EVALUATION WITH INSTRUCTION

Bayview Unified School District, an urban medium-sized low-wealth District serving a heterogeneous and growing population, has, within the last five years set a consistent direction for itself. Coming out of a decade of decentralization and local school autonomy, the District's orientation is now towards greater coordination of functions within the central office and the schools.

District personnel are unanimous in expressing a desire for better instruction to meet the needs of various types of students --those who are college-bound; vocation-oriented; basic-skills-oriented; need language instruction, etc. The new Superintendent, promoted from the number two spot last year, the Director of Instruction, the Coordinator of Staff Development, and the Coordinator of Curriculum constitute a leadership group within the District. They, with the support of the District Superintendent and some principals, have decided to continue a course of action that has a six-year history: to improve the skills of teachers through a combination of staff development courses and emphasis on principals' clinical supervision role. These individuals believe that teaching effectiveness using techniques such as task analysis and diagnostic/prescriptive sequences accompanied by other factors such as school and classroom climate and of team building will improve student learning.

Interest in testing and evaluation is relatively new within the District. There is no strong leadership or enthusiasm for using any particular testing program or evaluation activity as the organizer

for instructional improvement. There is general skepticism among the District's leadership group about the match between tests and evaluations and the District's instructional program, as well as fear about the community consequences of spotlighting low student scores. Nonetheless, compliance with externally mandated testing programs is causing the District to make productive uses of them.

Descriptions of activities in relation to the District's proficiency examination, the State Assessment Program, the Comprehensive Test of Basic Skills (CTBS), evaluations of Title I programs, evaluations of schools in the School Improvement Program will be detailed below. A short sketch of how the District is beginning to think about the relationships between testing or evaluation and instruction will provide a context for the later details.

Explorations are already under way in the following areas:

- Developing remedial procedures and programs for students not passing the District's proficiency exam during pre-12th-grade screenings.
- Developing staff development workshops and teacher-oriented materials aimed at assisting teachers in teaching the District proficiency skills.
- Matching high school departmental responsibilities for teaching with particular skills being tested by the District proficiency exam.
- Analyzing the match between State Assessment test specifications and District's curricular emphases.
- Using Districtwide student skill deficiencies identified by the State Assessment Program, CTBS and the District Proficiency tests as content for staff development course work in task analysis and diagnostic/prescriptive teaching.
- Planning for the development of a criterion-referenced testing system to back up newly coordinated Districtwide continua in reading and math.

- Making instructional inferences from test scores; e.g., CTBS scores showed that more instructional time combined with teacher support procedures increased learning by junior high school students using computer-assisted instruction.

In the next section, we will describe specific functions of the District as they relate to instruction. These include: testing and evaluation, staff development, continua, role of principal. (Not discussed because they appear, at the present time, to be marginal to the District's thinking about instructional improvement are materials; personnel resources such as aides, subs, volunteers; and budget support.) Following this discussion, we will describe the coordinating mechanisms which link these functions to one another and to the instructional program.

DESCRIPTION OF DISTRICT FUNCTIONS

Testing and evaluation. Most of the leadership group in the central office as well as teachers at Bayview do not find the assumptions underlying testing directly relevant to their view of teaching and learning. They believe education is personal and interpersonal. They would prefer developing the capabilities of teachers and students rather than restructuring instruction to conform to test specifications, that is, teaching to the test. Furthermore, they believe that tests do not provide an accurate measure of what students know and can do because of discrepancies between the content and skills tested by the tests and what the teacher teaches. Many of their students do not have test-taking skills and are not motivated to perform on tests. Nonetheless, they have demonstrated an openness to the possibilities

that examining test specifications and the patterns of student scores can lead to specific instructional adjustments.

PROFICIENCY TESTING by all districts in the state was mandated by the State Legislature in 1974. Each district was to develop its own examination. Each district was to develop a system for screening and providing remedial instruction for students before their last year in school. Students in the class of June 1981 who had not passed the examination would not be granted a diploma.

Professionals in Bayview as in other districts throughout the state are working towards compliance with the law. As educators, they are reluctant to deny students their diplomas. They also fear individual or class action lawsuits of the kind already initiated in Florida.

Bayview first developed their own tests related to curriculum content. Pilot tryouts revealed technical adequacies in the items. Community representatives also argued that the test should be built around real-life skills. A second version was developed in reading, writing and mathematics with the help of consultants who checked items for relevance to minimum life-coping skills and for sexist or cultural biases. Forms for grades 5, 9, 10 and 11 were also developed and have been used for two years. Fifty percent of the 8th graders taking the 1979 version of the examination did not pass.

The Coordinator of Curriculum says, "The proficiency exam has helped the District focus on curriculum. It has provided the impetus for reading, language and math continua and pointed out the need to develop Districtwide skills. For example, in math we teach computation but the test tests applications through story problems."

The Coordinator of Staff Development says, "We have to teach teachers effective teaching skills. We have to teach them how to teach the skills we are testing. Some teachers only lecture, make assignments, and send students to a lab. We want everyone to own this problem." She has organized a series of workshops to train the teachers who will conduct in-service in connection with the remedial summer school.

STATE ASSESSMENT PROGRAM testing has been mandated since 1974. The intention of these tests --one half hour per student on sampled items-- is to provide comparative data on how schools within districts are performing and on how districts within the state are performing. For several years, Bayview teachers and principals avoided facing the issues raised by the District's low performance and by the low scores of particular schools. The Director of Staff Development used an analogy to explain teacher avoidance. "It is like when kids get a D in a subject and teachers assume students don't know how to study or what to do. The teachers did not see the connection between instruction and the tests." She described the beginning of the change in attitude. "We (central office staff) tried to model a problem-solving way of looking at it so principals could do similarly in their schools. The Director of Instruction worked with the principals in the way he wanted them to work with teachers. Also, we asked teachers if they were addressing areas of the test. They said they were. When we observed, we found teachers had difficulty defining these skills to be taught as well as diagnosing for these skills. As a result, we built task analysis cycles into our Professional Development Center programs focusing on the low scoring skill areas identified by the State Assessment Program."

The School Board paid attention to the State Assessment scores. They wanted to know why the District scores were low, and asked for detailed analyses. Their interest, prodded by the newspapers who give widespread coverage to District scores, in turn prodded the District administrators to action.

The District administered CTBS TESTS in all Title I schools. This year, for the first time in a decade, the district required CTBS in all schools for grades 1-9. Teacher reactions range. "It's a waste of time." "It's useful to have to talk with parents." "Our kids cannot do it easily. They get upset when they cannot answer the questions."

The Coordinator of Special Projects who is responsible for planning, proposing, monitoring, and evaluating federal and state programs describes CTBS scores as primarily useful in preparing the needs assessment sections of subsequent Title I proposals and justifying programmatic activities.

Other testing and evaluation activities in the District relate to the bilingual programs. Whatever is done in this area seems to be compliance monitoring rather than program evaluation leading to instructional change. A similar position is maintained in relation to evaluating the statewide School Improvement Program funded in 5 of the 16 elementary schools.

Professional development. The current and central concern of the Bayview School District is instructional improvement through staff development.

The interest of the District toward professional development activities has its origin in several areas. The recently-departed Superintendent supported and encouraged team building and the school-site leadership.

In addition, he selected those principals and teachers active in this area by placing them in the central office in positions of leadership.

The principal and faculty of Hilltop School were such individuals. The principal, Mike Mitchell, in 1976, went to a conference where Madeline Hunter gave a presentation on team teaching and flexible scheduling. In addition, he joined a group of administrators in the Catalyst Program, a consortium started to explore innovative ideas for schools. Hilltop's staff began to experiment with a number of different organizational arrangements. Finally they decided to focus on the improvement of the instructional act. Mr. Mitchell invited Dr. Hunter to Bayview to give in-service workshops related to the Essential Elements of Instruction. The staff of Hilltop School worked together with Dr. Hunter and staff from University Elementary School at UCLA over a three-year period in an effort to improve the quality of teaching at Hilltop School.

The recently-departed Superintendent supported his staff in their decision to write grants in order to obtain federal money for staff development. Mike Mitchell, who was now the District's Director of Instruction, initiated the writing of a state grant that funded the first Professional Development Center (PDC) for the District.

Currently, $65,000 of the $110,000 used to run the PDC comes out of the budget of nine Title I schools. The purpose of the program "is to strengthen classroom instructional techniques in reading and mathematics in kindergarten and the elementary grades by involving school district personnel and student teachers from higher education institutions in pertinent professional development training programs. The

primary focus of this program is the continuing development of effective teaching strategies."

Augmenting the PDC program is the Teacher Center (TC) which serves seventeen schools including four parochial schools in Bayview. It receives a grant of $193,000 in federal money to implement the program. This program is intended to provide teachers with the opportunity to (1) develop and produce curricula designed to meet the educational needs of students; (2) use educational research findings or new or improved methods, practices and techniques in the development of the curricula; (3) familiarize teachers with developments in curriculum and educational research to improve teaching skills.

Both the PDC and TC emphasize teaching effectiveness skills for teachers. Under this rubric are:

1. Teaching to the objective
2. Ascertaining level of difficulty
3. Monitoring learning
4. Appropriately using principles of learning

Professional development in Bayview is teacher-oriented and principal-directed. Teachers are to learn a framework for teacher decisionmaking that is presumed appropriate for all subjects and grade levels. Additionally, principals are to learn how to be instructional leaders. All principals are required to attend workshops dealing with supervision. Principal training includes three levels: recognizing the elements of effective teaching; acknowledging and explaining effective teaching; and assisting teachers in improving their teaching.

Continua. Continua are documents which are intended to define the subject matter and skills to be acquired by students at given

grade levels. Usually, for teachers, continua compete with their own judgment, with the content and sequence of textbooks and with "normal operating procedures" in determining what content is presented to children.

In Bayview, during the decentralized '60s, each school developed its own scope and sequence. It appears justifiable to infer from the remarks of respondents that such continua were not controlling or binding on teachers. Teachers, within broad boundaries, did their own thing.

The District is moving towards greater coordination between the central office and the schools. There is continuing evidence of students' low scores on the statewide assessment tests though there is a definite trend upward throughout the District. The Coordinator of Staff Development observed that many teachers in Bayview did not use finely-tuned teaching strategies aimed at developing a common set of skills. These factors have stimulated the leadership group in the District to develop and implement Districtwide continua. Under the direction of the Coordinator of Curriculum, teacher committees have been meeting for more than a year. Their work, in the areas of reading, math, and language is in the schools, presumably being tried out before final revision and Board hearings about adoption next year.

It appears that the continua are soon to be refined documents consistent with either the State Assessment Program, the District Proficiency examinations, or the activities undertaken in the staff development area. Several District respondents stated their intentions to work in this direction; and to follow up such work with the

identification of benchmark skills ultimately to form the content of a District criterion-referenced testing system.

Role of principal. As former teachers and principals, the leadership group in the District office are acutely aware of the limitations of the District office in mandating changes in schools in light of the span of control and influence of the principal. Several former teachers from Hilltop have been appointed as principals in other schools. The Director of Instruction, himself the former principal of Hilltop, and the Coordinator of Staff Development believe that the principal's primary role should be that of instructional leader. Armed with the Madeline Hunter techniques for clinical supervision, the District is training principals for this role. The current Superintendent has "mandated" the presence of principals in the classrooms one day a week. He says that he, himself, spends at least eight hours a week in schools and classrooms.

We do not have sufficient evidence to determine the extent to which principals encourage their faculties to attend to the results of tests of any sort or to evaluations for the purposes of school or classroom instructional improvement. There appears to be a wide range of actions seemingly unconstrained by districtwide procedures. One principal said, "When our state assessment scores show us (our school) down in reading comprehension or skills, I tell the teachers they'd better do something - like go out and buy games and things to get those kids' scores up." Another principal reported, "I spend 20-30% of my time in instructional supervision. It's important for teachers to see me as a model. Sometimes I take the initiative; sometimes teachers ask me to come because they are having trouble in some area."

The Coordinator of Special Projects who handles Title I evaluations says that each school provides him with program data; he then combines it with student CTBS scores and returns to discuss his view of the findings and their implications with the principals and, on principals' request, with teachers. Whether this is a formality or an actual influence on instructional activities may depend on the personal relationships between the individuals involved, or on the philosophy of the principal about tests and evaluations. There does not appear to be a uniform, repeating, and predictable procedure for feeding back data to the schools and encouraging action based on the findings.

Summary. Bayview District has in place a functioning staff development program for principals and teachers. The District gives, because it is so required, the State Assessment Program tests, the District Proficiency tests and, for the first time this year to all students in grades 1-9, the CTBS. The District conducts, because it is so required, evaluations of Title I programs, bilingual programs, and other smaller federally-funded programs. At present, simply performing these functions appears to be a full-time job for District personnel. However, ideas about the desirability of linking these functions with one another are being discussed. Mentioned from time to time are other functions which might be brought into the instructional improvement effort: materials, definition of role of learning specialists, training for aides and substitutes, CRTs, etc. Simultaneously, ways of operationalizing such linkage are being imagined, argued about, tried out, discarded, or consolidated. The next section describes the linking mechanisms which are currently in place.

DESCRIPTION OF LINKING MECHANISMS

In Bayview, the informal and collegial relationships among the central office staff and between these individuals and school principals and teachers seem to outweigh the formal job descriptions, reporting arrangements, participation on committees, and attendance at meetings. Ideas about the potential linkages among testing, evaluation, instruction, staff development, and the supervisory role of the principal are beginning to emerge in conversations and discussions. Individuals' roles, responsibilities and reporting arrangements are shifting so as to promote coordination among these operations.

Within the District office there is a core leadership group. The Superintendent, in the manner as when he was the Deputy, shares in and supports the staff capacity-building orientation of the Director of Instruction and Coordinator of Staff Development, and the Coordinator of Curriculum, all former Hilltoppers. The Director of Special Services, who handles special education programs and their evaluations, and the Coordinator of Special Projects, who manages other federal and state programs and their evaluations, are influential but not central members of the group. The same is true of a newly-appointed Director of Research and Assessment whose role is as yet not clearly defined. The changes in staffing made by the new Superintendent in the past year have created anxiety for some individuals --especially those involved in testing and evaluation-- who are not certain of how their st-tus and role will evolve.

Morale within the central office appears to be good. In spite of the visibility of former Hilltoppers in positions of leadership, there is a widespread feeling that "people are interested in building bridges, not throwing stones." There is general acknowledgment that their

orientation towards staff development is sound, and that the leadership group are competent and hardworking. Furthermore, they are seen as demonstrating their own admonition to others to "model" desirable behaviors. Respondents cite approvingly their consideration of others' time and viewpoints and their attention to details and arrangements. Nonetheless, we heard the Coordinator of Special Projects and the Director of Research and Assessment express the wish that test scores and evaluation findings receive greater emphasis in the District. They felt some frustration with the lack of administrative support their views received.

While their reactions may have been influenced by their peripheral status within the informal hierarchy of the central office, they expressed their concern in terms of specific aspects of the current situation --admittedly still in flux-- such as access to the Superintendent and to superiors, geographic proximity to colleagues (the offices of the District are in two separate buildings); inclusion in decisionmaking committees or meetings.

Because of the central office's staff development emphasis, there is frequent contact between administrative staff, principals and teachers. In addition, teachers are represented on the District Proficiency Exam Committee. This group is concerned with developing remedial instruction for students who don't pass the screening examinations. Teachers are also represented in the groups forming the District continua.

Summary. For the present, then, the linking mechanisms in Bayview can be described as follows: There appears to be a high level of consensus among central office staff on the value of improving teacher skills

and stimulating principals' clinical supervision role as the way to better meeting the needs of the diverse student populations. There are several individuals who would like to emphasize test scores and evaluation results as input into staff development and principal monitoring, but their emphasis is not as yet a dominant one.

IMPACT ON CLASSROOM PRACTICES

The following comments are based on interviews with nine administrators and eleven teachers at five Bayview schools - 3 elementary, 1 junior high and 1 senior high. Interviewees were preselected according to their availability and their willingness to be interviewed. They were asked questions about what they did differently in their classroom practices now as compared with several years ago, and how they were influenced by various activities related to testing, evaluation, staff development, district continua, etc.

Without exception, teachers were aware of the District's Professional Development and Teacher Center programs although not all were clear about the differences between the two, and the current form of organization of the courses. Almost everyone had participated in at least one session of the PDC. In general, they were supportive of the ideas of the PDC although some said they had initially been skeptical. Older teachers said they had not encountered the content of the Hunter system in their teacher preparation courses. For others, it was a review appreciated after they had acquired some classroom experience. One said, "I'd rather have learned about instructional strategies, or about Magic Circles than about task analysis." Others said the elementary and secondary teachers should attend different sessions. Another suggested special courses for first-year teachers.

Awareness of the content of the courses appeared to persist for some teachers into their classroom teaching. They "thought about" objectives, diagnosing children's needs and sequencing instruction. It was difficult to determine whether what they did changed because of such awareness.

The District continuum, probably because it was distributed and used

on a trial basis did not appear to exert much influence over teachers' selection of content to teach or skills to ensure that their students mastered.

The District Proficiency Examination and the concomitant discussion about remedial procedures for nonpassing students were of concern to teachers. Most saw the remedial efforts as necessary and important; such efforts were described as isolated activities, however, rather than as a first step to rethinking either curriculum or instruction.

Summary. In Bayview, as in most other districts, there is great diversity in educational philosophy and practice among and within schools. The PDC training has not reduced the diversity in classroom practice. It has made a beginning in giving teachers and principals a common language in which to talk about teaching. The PDC Program also provided teachers with a shared professional experience from which further interactions may come. At the present time, Bayview's teachers, also in common with teachers in other districts, range in their attitudes towards testing and evaluation from hostility to indifference to an openness regarding analyzing test data for instructional implications. Among the mandated tests administered in Bayview's classrooms, only the proficiency exam seems to have immediate potential for impacting instructional practice.

NORTHTOWN SCHOOL DISTRICT
1980

BACKGROUND INFORMATION

Northtown School District serves a large metropolitan area, and enrolls students who are diverse in socioeconomic and racial characteristics. Since the 1950s, the population has changed dramatically from a primarily white majority to an ever-increasing percentage of Blacks, Hispanics and Asians.

Overall, the District is experiencing declining enrollments. However, because of population shifts and geographic constraints, it is in the unusual position of closing down some central city schools while building new schools in recently developed outlying areas.

The District operates close to 200 schools, K-12, and employs over 4,000 teachers. The School District is presently under court order to desegregate its schools and is facing possible court-ordered busing if appropriate steps are not taken to ease the reality and effects of racial isolation. A major concern for the court, the community, and the School District is pupil achievement on the CTBS battery.

The District has searched nationally for its teaching and administrative staff and has been able to maintain high staff stability over the last 20 years.

Beginning in the early 1960s, the District adopted a decentralized management approach. Accordingly, considerable authority for instructional and curricular development and practice rests at the local school site. As a result, there is considerable variation in the instructional program from school to school. The District uses the State Assessment Tests, the Comprehensive Test of Basic Skills (CTBS), and is developing a battery of criterion-referenced tests (CRTs). Because of the widespread community and court interest in CTBS scores as a measure of the District's success, the District has sought ways to integrate the CTBS tests into its decentralized instructional and curricular decisionmaking structure.

At a large number of District schools which have special funded programs, the District is required to give norm-referenced tests each year to every child. In these schools the District evaluation office has devised an individualized system that aggregates CTBS scores by school and presents them in a way that maximizes their use for school-site decisionmaking. The District maintains an extensive in-service training department upon which school sites can draw to help them implement their instructional programs.

RECENT HISTORY OF NORTHTOWN

In order to understand Northtown's current evaluation and testing programs and teaching activities one must know the District's recent history because, in part, the present program is a natural extension of previous decisions and structures.

Immediately following World War II, Northtown experienced an explosive population growth that placed enormous pressures on the then relatively small school district. The District could scarcely build schools rapidly enough to keep up with increasing enrollments. Because so much attention and resources were focused on buildings and facilities, there was insufficient parallel change and growth in the District's organization and instructional programs. Soon it became apparent that the programs of the '40s and the '50s would not be adequate for the decades ahead; a new organizational structure and educational program was needed.

One of the pivotal figures in determining the District's new directions was Dr. Wilcox, who served as a Deputy Superintendent for many years until his retirement in the late 1970s. As Research Director in the mid-1950s, he instituted a new testing procedure, and his assistant (who subsequently became Research Director) installed an IBM card system for more rapid scoring and reporting of test data.

As Assistant Superintendent, Wilcox exerted considerable influence over subsequent District developments. He served as the District liaison to an outside consulting firm that in 1962 was brought in to devise a new District management plan. He subsequently had a major respcnsibility for implementing and managing many of the consulting firm's recommendations.

Among the recommendations and ideas that came from a variety of sources, and in which Wilcox strongly believed, were:

1. A District's effectiveness is largely determined by the quality of the people it hires. Beginning with the period of rapid growth and continuing on to today, the District conducts nationwide searches for the best people available. Many of the District's present staff, from Superintendent through the various administrative and teaching ranks, were recruited from leading universities and school districts nationwide.

2. Parents and citizens should be actively involved in District and school-site decisionmaking. Parents and citizens are represented on a large number of Districtwide and school-level committees.

3. Authority should be decentralized as much as possible, to the local school-site level. The central office can encourage change and provide the means for change (in-service training, materials, aides, etc.) but the local school site must play a major role in identifying its own goals, determining plans to accomplish them, and implementing those plans. Wilcox tried to spend one day per week in the schools seeing to it that the schools responded to these responsibilities.

One of the results of this "decentralized" system has been a considerable proliferation of District instructional programs. The notion of a single Districtwide curriculum became outdated in Northtown. This proliferation was considered desirable because no single unified curriculum could possibly respond to the needs of an ever more diversified pupil population. Each year the Board adopts the Course of Study for Elementary Schools and the Digest of the Secondary School Curriculum. Moreover, within this framework, school staffs are encouraged to:

> "further develop and particularize goals, objectives, and priorities to meet the needs of the school, department, class, and individual student..."

As the District became increasingly diverse and complex it became necessary to design procedures that would bring some sense of order and facilitate communications. To that end, an elaborate integrating committee structure was formed. This structure was designed to make sure that representatives of those who will be affected by District-level decisions would have a chance to influence those decisions, know the basis for the decisions, and be informed about them.

The main committees are: Curriculum-Instruction Committee, Special Activities Committee, School Renewal Committee, New Programs Committee, and Superintendent's Leadership Council. They are designed to perform specific screening, advising, decisionmaking, and development functions; memberships cut across functional lines. The key integrating committee is the Curriculum-Instruction Committee. Its membership, to illustrate its cross-functional design, consists of: the Assistant Superintendent for Instruction, and one or more representatives from the Elementary Schools Department, Secondary Schools Department, Personnel Office, Student Services Division, Community Relations Office, Business Office, Evaluation Office, and teachers, students, principals and citizens. The Superintendent's Office provides a procedures consultant.

This Council monitors processes for instructional program development, reviews all proposals for program changes, and makes recommendations to the Superintendent's Leadership Committee. Interestingly, under this system much of the impetus for change comes from below. Proposals from all levels across the District are considered by this group which is composed of individuals who represent various constituencies who might be influenced or affected by proposed program changes.

Beginning in the mid-1960s, when this heady period of expansion and proliferation was at its zenith, the District began to undergo a fundamental and far-reaching change. The minority populations began to grow rapidly. First Hispanic and Black populations, and more recently Pan Asian populations, grew in their respective percentages of the District's overall population.

It was thought that the decentralized, school-based organizational and functional structure that had been developed would be the most effective way to meet the needs of an increasingly divergent student population. Over the years certain schools identified as Racially Isolated Minority Schools (RIMS) have had unusually high enrollments from one or more minority populations. Student achievement district-wide, as measured by norm-referenced tests, has been quite high, especially in comparison with other urban districts. This gave credence to the belief that the centrally supported, but essentially decentralized organization was working. However, in 20 RIMS schools it has become increasingly evident that there was great disparity between their pupils' achievement and the achievement of pupils in the other District schools.

This disparity developed in spite of the fact that the District was participating in numerous federal and state compensatory programs designed to improve these schools' performance. In the mid-1970s, the District received court-ordered desegregation. In response, the District has initiated a number of programs and activities to improve the educational programs and pupil performance in the Racially Isolated Minority Schools. These programs and activities include: Bilingual programs, Learning Centers, Alternative Schools, and Career Centers.

Recent CTBS results indicate that these efforts have stabilized the District's achievement scores but in the RIMS schools the hoped-for improvement of pupil achievement has not materialized. The scores remain low - the public, especially minority parents, are skeptical and untrusting of the District's efforts and motives. The use of forced busing is increasingly being considered by the court as a necessary next step.

The court and the Integration Task Force have increasingly pressed the District to take further steps that will improve these schools' CTBS performance. The District administration has been searching everywhere for some system or plan that will work. The District administration has increasingly limited the freedom of Racially Isolated Minority School staffs to try to solve these difficult problems individually. The District has selected and implemented several ideas or programs. Among the ideas and programs selected are:

1. Time on task. The District's evaluation unit played an important role in the Beginning Teacher Evaluation Study (BTES) research on time on task. Believing that time on task is a critical variable that explains student learning, the District has instituted stringent procedures to reduce classroom and school interruptions so that teachers can increase classroom academic learning time.

2. Adoption of specific teaching programs. The District has searched for specific, carefully developed teaching programs that have shown evidence elsewhere of having achieved results in schools similar to the District's Racially Isolated Minority Schools. One such program, Distar, is a very structured and controlled reading instruction program.

Some members of the central administration are impressed with Distar's results elsewhere; they think it is worth a try. Distar is being implemented in selected classrooms in Racially Isolated Minority Schools.

3. Mastery learning. Based on Benjamin Bloom's work at the University of Chicago, and his research in the Chicago and other school districts, Northtown is instituting, in the fall of 1980, a mastery learning program in the Racially Isolated Minority Schools.

In Spring 1980, the District, in essence, signed an agreement with the court that they will guarantee that the Racially Isolated Minority School pupils' CTBS scores will rise by 9 percentile points over a three-year period. At the beginning of the 1980 fall school year, the judge has indicated that even this may not be enough --the School District may have to guarantee even higher performances. District failure to reach these goals may result in court-ordered busing.

The result of all this is to place the District at a crossroad. It is essentially trying to maintain one kind of District plan and structure for the majority of its schools (decentralized, relatively high local autonomy) and another structure for its Racially Isolated Minority Schools (centrally prescribed, highly structured programs, with guaranteed results).

Such a course is fraught with problems. Among them is trying to convince teachers and principals in Racially Isolated Minority Schools who are accustomed to autonomy, that this new, more structured approach will achieve better results and that they, the principals and teachers, should thereby give up some of the relative autonomy they enjoyed previously. Thus, as the 1980 year begins, the District is in a very

precarious position. It has been forced to adopt a more centralized, controlled management system in some schools, but not in others. The District itself has guaranteed results in some schools while not knowing if it has the technical knowledge and delivery system to make good on their promised performance.

Likely, pupils' performance on CTBS will have a profound effect on this development. That is, if mastery learning and specific structured instructional programs and time-on-task do indeed "turn these schools around," this will justify the District's program. Perhaps then the Racially Isolated Minority Schools' teachers and administrators will be more tolerant of their somewhat impaired freedom. If the scores do not go up as hoped and if forced busing results, the District may face even more severe problems in the years ahead.

Testing and Evaluation Program

Given this brief historical sketch, we turn now to a discussion of the development and present functioning of the District's evaluation and testing programs.

As with most urban districts, Northtown's evaluation and testing activities have developed largely in response to state and federally-mandated evaluations of funded programs (e.g., ESEA Title I, ESEA Title II, School Improvement Program, ESEA Title IVc, State Compensatory. In the 1979-80 school year the District reported conducting evaluations of many different projects that were funded at the federal, state, or local level. Many of the evaluations are conducted by outside evaluators (e.g., ESAA, ESEA, Title IVc, Follow Through). The main internally-evaluated programs

include: ESEA-Title I, Pre-school programs, Reading Demonstration, the local Master Plan, and GATE.

Each year the Evaluation Services Office submits an Annual Test Report to the Board of Education. This report typically includes statewide testing and CTBS results. In addition, the Evaluation Services Office conducts special nonmandated evaluation and research studies. Often these studies are requested by District level administrators regarding some ongoing District activity or program, or about some proposed program. A recent example was a special report on the BTES Interruption Study which led to a District policy to reduce interruptions and thereby increase time on task in Racially Isolated Minority Schools.

The two main testing programs administered by the Evaluation Services Office are the Districtwide CTBS and the required state test battery. CTBS is given Districtwide in grades 5, 8, and 11, and is given annually to all students enrolled in special project schools, e.g., Title I. The scores are used differentially by various groups. The Board, the Citizens Integration Council and the court are particularly interested in CTBS score gains in the Racially Isolated Minority Schools. The principals and the compensatory education staff use CTBS in writing SIP and Title I reports and in program planning. The District staff uses CTBS results as one means of checking on Districtwide instructional programs and providing necessary remediation. For example, they discovered that they were not teaching the application of fractions, which was tested by CTBS. Subsequently, the instructional program was revised and teacher workshops and special materials were provided.

The State Assessment tests are administered to students in the 3rd, 6th and 12th grades. An annual report of the results is submitted to the Board. These test results are analyzed to see if they reveal instructional or curricular deficiencies. The test is less onerous to teachers than some because it doesn't take up too much of the students' time. On the other hand, the teachers seldom referred to the test results as having any influence on their teaching methods.

Interestingly, the proficiency testing programs (CRTs) are almost totally isolated from the Evaluation Services Office. CRTs are considered part of the District's curriculum program, and the District's curriculum staff develops, administers, and interprets the CRTs. A member of the Evaluation Services staff sits on the CRT committee but he reports having little influence.

It appears that of the tests the District administers, the CTBS program currently has the greatest impact on the District's decision-making. This is largely due to the wide public interest in the Racially Isolated Minority Schools and the way CTBS scores have come to be considered the primary indication of the District's sincerity, competency and commitment.

The Evaluation Services Office is keenly aware of the limitations of CTBS scores as an accurate measure of a school program's effectiveness. They know about the many factors that can influence CTBS scores, and the need for caution in using CTBS in this way. Given the possibility of Districtwide busing, however, they must comply with this CTBS "fixation." The evaluation staff is making the best of what it considers an untenable position.

Because the District has emphasized individual school-site development of instructional programs (within limits) the District has not made any systematic effort (until recently in the Racially Isolated Minority Schools) to coordinate testing, curriculum and evaluation in any single, Districtwide manner. Instead, the District has been developing a testing and evaluation reporting system that is geared to the needs of each individual school; this system is consistent with the District's long-held belief in local school-site autonomy. Limited presently to its consolidated application* schools, the process can be briefly described as follows:

Each consolidated-application school's CTBS scores were obtained initially and presented to each school's principal and staff along with the school's mobility index, minority percent, and school enrollment figures.

Based on these data, the school staff, with the assistance of an Evaluation Services Office evaluator, determines a set of objectives and activities for the coming year. These form the core of the school's annual improvement plan. District evaluators regularly revisit these schools during the ensuing year. The CTBS tests are administered again in the Spring and individual pupil results are reported to the appropriate teacher before the end of the school year. During the Summer, the Evaluation Services Office staff scores the tests and analyzes the results in terms of the individual school's stated goals. A school-specific report is prepared and presented to the school staff in the beginning of the Fall quarter. This forms the basis for the school staff to reformulate goals and activities for the next year--and the cycle is repeated.

*The state has developed a common form (Consolidated Application form) so that districts can provide basic demographic data once while applying for several state and federally funded programs.

Every attempt is made to individualize each report for the specific schools and to report the data in an easily readable style, e.g., use of histograms. It is important to note that this procedure is consistent with the District's belief that the instructional decisions should, as much as possible, be made at the school-site level.

Supporting these school-level efforts is a very large District in-service training program which mandates a limited number of activities in support of the desegregation efforts (e.g., human relations workshops) and provides a wide array of other activities from which individual schools, principals, or teachers can select. Presumably schools, on the basis of their annual plan, will be provided the in-service necessary to implement their plan.

Impact on Classroom Practices

Based on relatively short interviews with a small sample of teachers, the District's testing and evaluation program's impact on classroom teaching practices seems to be quite minimal. The tests that seem to have the greatest impact are the District CRTs, especially those used in conjunction with the state-mandated proficiency testing program. Often this primarily influenced those who teach grades and subjects that are affected by the state proficiency program.

Teacher responses differed, of course, but in general teachers were aware of the CTBS program and believed the scores probably meant something to the central administrative staff who had to answer the Board's and community's questions about pupil performance. The teachers did not refer to CTBS or SAP when discussing those activities that influenced their teaching behavior.

The activities reported as being more helpful and thereby having the greatest impact were some of the District's in-servide training programs and the aides and resource specialists available to them. Many teachers did report changing their behavior over the years --but testing and evaluation were seldom mentioned as reasons for those changes.

In the consolidated application schools in which we conducted interviews, practically no mention was made of the District's school-site individualized evaluation program (described in the previous section). This lack of knowledge or response could be a function of many things, especially the newness of the program. It may be that, as the school staff gain experience in using tests this way, this program will increasingly penetrate the local school's decisionmaking process. It may be, also, that our measures of impact were insufficiently sensitive or comprehensive. On the basis of these initial interviews, however, we heard little that indicated those special evaluation reports were presently having the desired effect at the local school-site level.

Indeed, teachers most often reported that the various testing programs consumed far too much classroom time, and, in some instances, imposed an inhumane burden on many kids who had to take an inordinate number of meaningless tests. Several doubted the accuracy of many test results because the students, who are "over-tested," simply burn out and do not try on the tests. This results in school-level testing profiles which are inaccurate reflections of the school's work and progress.

SHELTER GROVE SCHOOL DISTRICT
1980

BACKGROUND INFORMATION

Shelter Grove is a small elementary school district consisting of seven elementary schools in which there are 132 teachers, 7 learning specialists, 7 principals and 3,000 children.

The community of Shelter Grove is relatively homogeneous. Only 10% of the children going to Shelter Grove schools are minority. Very few of them have language problems. The adults in Shelter Grove are mostly professionals or work in technical occupations. Shelter Grove is a bedroom community serving a variety of urban centers located within 50 miles of the community.

The community has been stable with very few people moving out. The population has been gradually increasing due to new housing in the area.

The District is likewise stable. Fifty-five percent of teachers have been in the District more than ten years; 46 percent of principals are long termers. There is a small central office consisting of 5 professionals and 15 support staff. Eighty percent of these individuals have been with the District more than ten years.

The District has called itself "poorer than average elementary district" citing that, in 1974, it was at the 21st percentile in dollar expenditures per pupil as compared with other California school districts. It has averaged around the 31st percentile over the last decade.

Changes are likely in the District due to three factors. The Superintendent and Assistant Superintendent are departing as of September 1980. Concurrently, the Shelter Grove Elementary School District will become a unified school district; five of the elementary schools will remain as is, two will become middle schools, and an existing high school will become part of the District. Sixty teachers will be moved. A third factor influencing the future of the District is funding. Because of the presence of military children, the school has received federal PL 875 funds instead of tax dollars. Anticipated changes in the law governing these funds and anticipated changes in state funds for schools may require budgetary revisions within the District.

WHAT SHELTER GROVE IS DOING TO LINK TESTING AND EVALUATION
WITH INSTRUCTION

Shelter Grove seems an ideal community in which to try an educational experiment leading to improved educational excellence for children. The community and the school district seem not to have been beset by many of

the social problems plaguing other areas in the country. There has been no major increase or decrease in population. There exists no large group of children with English-language difficulty. There exist no major political or economic divisions within the community.

Interviews with Board members, central office personnel, principals, learning specialists and teachers revealed remarkably homogeneous perceptions about instructional purposes. The president of the school board said, "Almost everyone believes in and works hard at teaching individual kids. The kid is the most important thing. We try not to have any throwaways." A teacher said, echoing the sentiments of a dozen of her colleagues, "This District expects a lot from its teachers; it's a great place for kids; they really learn. I moved here so I could send my own kids to the schools in this District." The coordinator of materials said, "We really concentrate on having the children learn --basic skills first, as well as the other important things. The parents would not have it any other way."

Within the District there seems to be general consensus that learning is important and that children are important. Although this is the rhetoric of most school districts, professionals in Shelter Grove seem to be willing to act in light of this concern even when such actions require more work, some reorientation in their thinking, some readjustment of territories.

The connections between testing or evaluation --that is, data collection enterprises-- and instruction are seen differently by different people in the District. The Superintendent has a management orientation. He sees testing and evaluation as inextricably connected to instruction. "Testing and instruction are intimately related." The Assistant Superintendent is curriculum-oriented. Her experience working with curriculum and working with teachers perhaps has influenced her twofold emphasis on staff development

and materials and equipment. The principals in Shelter Grove see their roles as instructional leaders. They are required by the District to spend time in each classroom. They are expected to be intimately familiar with the instructional program as well as the progress of individual children within their relatively small schools. The teachers see themselves as hard-working and dedicated professionals.

The current management of instruction is closely linked to the following activities: testing, a District continuum, a variety of materials mediated by a media specialist. These instructionally-linked functions are supported by the Professional Development Program for principals, teachers, aides and substitutes; by learning specialists in each school; by the conceptualization of the principal as evaluator and facilitator of instruction. Furthermore, the instructional management is aided by recruitment and selection procedures for new staff as well as budgetary considerations. A description of each of these functions is presented in the next section to be followed by a description of the coordinating mechanisms which link them to one another and to the instructional program.

DESCRIPTION OF DISTRICT FUNCTIONS

Testing and evaluation. The District administers a number of tests. The COMPREHENSIVE TEST OF BASIC SKILLS (CTBS) is given annually to the 200 children in the two elementary schools participating in the Title I program. Within the District, CTBS scores are not regarded as useful or indicative of student learning. CTBS tests are used primarily to comply with the evaluation requirements of Title I.

The District administers to grades 1, 3, and 6 the State Assessment Tests in conformity with state regulations. These tests are designed to

compare schools within the District to one another and to compare the District scores in reading, writing, mathematics, spelling, and other language arts with similar districts in the state. The Test Director of Shelter Grove finds that scores from the State Assessment Tests are useful in the following ways:

- to use in public relations with the media and the parents, since Shelter Grove children generally score quite high.
- to examine, in a gross way, the performance of children in certain subject areas. For example, spelling scores dipped one year and so attention was paid, through staff development courses and curricular reexamination, to improving spelling skills of children.
- to examine long-term trends in the District. However, he notes, that yearly fluctuations are to be expected, since there is a natural tendency for the scores to "regress towards the mean," especially since Shelter Grove children score at the high end of the normal curve.

The STATE ASSESSMENT TESTS are not highly valued by others within Shelter Grove; neither are they a major encumbrance. Teachers note that the test takes only one half hour to give, that the students do not seem to mind answering the questions, and that they have some curiosity about the results but no anxiety connected with the scores.

All seven elementary schools in Shelter Grove participate in the statewide SCHOOL IMPROVEMENT PROGRAM. The schoolwide planning and the evaluations --on site and conducted by a three-member team trained and organized by the state-- seem quite compatible with other District efforts. "We were doing this before the School Improvement Program came along."

The evaluations connected with the School Improvement Program seem to be quite salient for the principals and teachers in Shelter Grove. This appears to be, in part, accounted for by the personal nature of the evaluation. Teachers or principals from neighboring districts are recruited by the state to serve as members of an evaluation team. They are quite familiar with the program at Shelter Grove and are seen as informed and interested people. In essence, it is evaluation through personal judgment.

Since the quality of these evaluators is now good in the opinion of Shelter Grove teachers and principals, there is a good deal of attention paid before the evaluation and during the site visit to having things in order. The District, for example, runs mock site visits using teachers from one school who visit other schools in the District. There is substantial competition among principals to have their schools score the maximum. An anecdote very much on the mind of District personnel concerns the recent visit of an evaluation team. The team completed the evaluation of one school and gave it the highest ratings that could be earned. There was some concern among central office personnel that the principal of the second school would be enormously upset if his building did not perform as well. Fortunately, the second school also scored the highest possible rating a day later. There was, apparently, a great sigh of relief throughout the District.

The District itself uses an annual ATTITUDE SURVEY of students, of teachers and of parents to ascertain their degree of satisfaction with the program. The student questionnaire asks self-report items relating to the child's perception of himself or herself as student in particular subject areas as well as his or her feelings about the school, the classroom and the teacher. The adult questionnaires ask for opinions about the functioning of the school program. This information is analyzed by the Testing

Director who reports it back to the principals and teachers on an annual basis.

However, the foregoing tests and evaluation procedures do not have the power to affect instruction in the same way as the District's CRITERION-REFERENCED TESTING SYSTEM. This system, developed over time by teachers, is the major device regulating instruction. The CRT tests, each taking no more than half an hour to administer, are given three times a year or more often, at teachers' discretion. Normally, teachers teach a unit of material in math, reading, or language arts, then wait two weeks and test their students. The test booklets are scored by the teachers and sent on to the Testing Coordinator who returns to them printouts with their students' scores, organized by objectives, printed out by learning groups.

The use of CRT scores in Shelter Grove can be described most simply according to the school calendar. In May of each year the scores are aggregated by the Testing Coordinator into profiles for individual students and for instructional groups. These scores are made available to the schools prior to their two planning days. It is during the first planning day that the program of the school is evaluated by looking at the group, classroom, and total school scores. The principal conducts this planning session and regards it as a needs assessment for the following year. The Testing Director is available for assistance in interpreting the test scores. On the second planning day toward the end of school, teachers set up their next year's learning groups using the profiles that have been assembled on each student. In September, prior to the first day of school, the test information is updated. This updating involves changing programs from one instructional group to the other depending upon teacher assignments, and adding new students who have come into the District. Within the first week of school teachers get profiles of their students by their new instructional groups.

In October, teachers confer with parents using the objectives printed out for the criterion-referenced tests. They indicate to the parents what the instructional objectives are for the children during the coming year. Also in October the teachers meet with the principals and set the goals for the children in each instructional group.

Between October and February there are two or three data collection points, that is, the CRTs are picked up, scored, and the information returned to the teachers. When the teachers in each building get their printouts, they meet with learning specialists in each school to plan any revisions which appear necessary in the instructional prescription for particular children.

In February there is a midyear evaluation. Summary reports on each child, each group, each classroom and each school are run. The principals meet both with the teachers and with the Superintendent and the Assistant Superintendent about the progress that students are making. Plans for modifying the instructional program are made.

In April and May proficiency tests composed of various segments of the criterion-referenced tests are administered to children in grades 4, 5 and 6 in accordance with state requirements. In May letters are sent to parents for the several children who are performing at two grade levels behind. Conferences with parents explore problems as well as teachers' intentions for improving the learning of those students.

In June, the two planning days previous described occur.

The criterion-referenced testing system in Shelter Grove is coordinated with the District continuum in math, reading, and language arts.

District continuum. The District continuum appears, from the report of teachers, to guide their selection of materials to teach to students.

It appears to be more influential in the areas of reading and language arts than in math. In mathematics, some teachers report that they use the textbooks as the sequence of instruction checking that the District continuum objectives have been accomplished. The continua are written in looseleaf books which appear to be well thumbed and readily accessible to all teachers. The District continua grew out of attempts to coordinate an earlier version of texts with goals and objectives. The continua as it evolved resulted in the replacement of the earlier CRT system with test items that matched benchmark objectives.

Materials. Within every school in Shelter Grove is a materials and media center staffed by a full or part-time specialist. The District orders materials that teachers and learning specialists request. The central office coordinator of materials each year asks the teachers what they want, look at areas in which materials need to be supplemented, and meets with the media specialist to facilitate the acquisition of equipment and supplementary curriculum materials. The learning specialist in each school often works closely with the media specialist to try out new materials as they come in to determine the extent to which they are appropriate for particular types of children. The learning specialist may try out materials sent to the schools as samples or on a tryout basis to determine whether they are suitable for ordering. The media specialist in each school often functions in an advisory capacity to the learning specialist and to the teachers, recommending to them books, films, adjunct materials which he/she thinks appropriate.

Professional Development Program. The Assistant Superintendent believes that the PDP is at the very core of the support system which links testing with instruction. The PDP intends to provide teachers with the skills necessary to act upon the results of the CRT tests.

The PDP is coordinated by a full-time specialist. The program has evolved over time. Currently it includes a level one which is based upon the Madeline Hunter orientation to teaching. All new teachers are required to participate in level one activities. A modified version of level one activities is provided for substitutes and aides. The schedule for level one activities includes three sessions, several weeks apart. The first session deals with goals and objectives, designing a lesson, giving effective directions. The second session deals with motivation and reinforcement. The third session deals with task analysis and diagnosis. These sessions are scheduled several weeks apart. The sessions cover content which can be applied in the classroom. The teacher participating in level one PDP activities invites the coordinator into his/her classroom to observe an application of the session's instruction. This happens on three occasions.

Older teachers are encouraged to go to the Professional Development Center sessions and most (70%) have. The teachers are relieved of classroom responsibilities by qualified substitutes. They are released for fifteen half-days for the level one program, they are released for eight half-days for the level two program.

Level two programs are of two kinds. The first is an extension of the level one activities with attention to individualizing instruction by choosing among alternatives. Additional sessions cover matters such as diagnosing and developing teacher strategies for meeting affective needs, using inquiry skills techniques, teaching to specific curriculum content such as reading, spelling or writing.

Prior to the widespread implementation of the Professional Development Program, the principals were required to take the level one course as well as additional courses in clinical supervision.

The widespread participation by principals, teachers, substitutes, and aides in the level one PDP activities has had some interesting consequences.

- Teachers and principals as well as learning specialists, aides and substitutes all now have a common language to talk to one another in. Clinical supervision by principals is made easier by a common frame of reference and a common understanding about teaching strategies.
- "All teachers use the same strategy: styles vary considerably," was the way the coordinator of materials described the coherence among teachers. The PDP appears to set a broad framework around instructional activities within the classroom.
- The tools of teaching provided by the PDP level one courses are consistent with the continuum and CRT testing system. These tools include attention to objectives, teaching strategies, use of data, individualized instruction. Teachers and learning specialists therefore have methods for acting upon the results of the CRT tests and making prescription for the improved learning of particular children.
- Level two courses and supplementary programs are optional. Teachers, however, express interest in participating. In part this is because they get release time from their classes and subs who know the children are available to take over. In part there is a general orientation in the District that professional growth is valued; professional growth is presumed to be facilitated by courses offered through the PDP. Many teachers feel that this is in fact so. There are some who would rather remain nonparticipants.

Learning Specialists. Within the last three years the role of learning specialist has evolved within the school district. The learning specialist is now regarded as a master teacher, not as an administrator. When the position was first conceptualized, it was not viewed positively by teachers. The position has now become institutionalized so that teachers see the learning specialist as helpful to them and available to solve their problems.

Approximately 40% of the time of the learning specialist is spent working directly with children. These children are those identified by the teachers and by the specialist together as in need of special remedial help. The other 60% of his/her time is devoted to two kinds of activities: first the learning specialist works with teachers in their classrooms to provide instruction meeting the needs of children as indicated by the CRTs and by teacher observation. The learning specialist sometimes conducts demonstration lessons. The learning specialist sometimes makes suggestions for new materials. The learning specialist meets with teachers when the CRT scores come back and helps plan for the next several months. The small remaining part of the learning specialist's time is devoted to in-school-in-service activities. For example, he or she presents changes in the continuum or new policy directions from the District to teachers within the schools. The learning specialist is responsible for diagnosing new students as they come into the classroom and facilitating the placement of children in instructional groups, especially if those instructional groups cut across classroom lines.

In the opinion of the professional development coordinator, the principals, many teachers, the coordinator of materials, the role of the learning specialist is critical to the linking of testing and instruction in the District. The learning specialist is a resource to teachers and makes

possible acting upon the prescriptions implied by children's responses on CRT tests. The learning specialist brings to the teachers suggestions and possibilities for instructional alternatives; more importantly, he or she provides extra instructional time for children who need it. Essential to the role of learning specialist is its separation from administration and evaluation. The services provided by the learning specialist are facilitative not prescriptive. He/she is seen as a resource for teachers not as an intruder in the classroom.

Principals as instructional leaders. Principals are expected in this District to spend a good deal of time in the classrooms. Teachers report that principals are invited in to watch lessons with regularity. Following each of the observations the principal and the teacher confer using the concept and language imparted in the Professional Development Program. Principals have an evaluation function, "I am the clout that backs up the learning specialist," says one principal. I evaluate, the learning specialist helps."

The District administration expects each principal to know intimately what is going on in the classrooms. Such topics are formally discussed at a February meeting with the Superintendent and Assistant Superintendent. Informally, they are discussed at weekly administrative council meetings.

Recruitment and selection. Further removed from the classroom than the foregoing functions, but nonetheless contributing to the instructional program, is the recruitment and selection of new teachers and new principals. Teachers are fit upon the committee to hire new teachers. Principals sit upon the committee to hire new principals. For teachers, in addition to the usual interviewing procedures, there is a requirement that a demonstration lesson be presented with a small group of students. This demonstration

lesson indicates to the observing teachers the extent to which the applicant already has the philosophy and the skills which pervade teaching throughout the District.

DESCRIPTION OF LINKING MECHANISMS

In this section, we will describe the organizational mechanisms that bring together three elements of the instructional program --classroom practices, district continuum and criterion-referenced testing. We will similarly outline the linkage among functions that support the instructional program --staff development, materials, the roles of the learning specialist and principal. Finally, we will describe how personnel and budget functions relate to instruction.

But first, it is essential to bring some of the intangible preconditions of such linkages into view. In Shelter Grove, there seems to be, at this time,

- relative homogeneity of purpose. There is expressed among all levels of respondents a commitment to a teaching-testing-teaching orientation to instruction. The Superintendent says, "Central office staff have worked together for so long, they have similar attitudes." The Board president indicated strong parental support for the technical orientation to teaching characteristic of the District's approach. Teachers who were initially opposed to this approach have either left or been socialized into modifying their position, at least publicly.
- recognized high level of skill. Central office staff say about one another, "He is very good at his job." Learning specialists give anecdotal evidence of how their principals do their jobs. The staff development specialist talks about support from the Assistant Superintendent. Teachers have high regard for the staff development specialist, for the testing director, for the learning specialist.

- generally supportive atmosphere combined with task-oriented competition. The seven principals work well together although there is some evidence of an inner leadership group. The Director of Personnel says the principals "have a great deal of influence in the system, after the Superintendent and the Assistant Superintendent." The Coordinator of Materials and the Assistant Superintendent are old friends as well as colleagues. The Test Director has moved up within the system, having been on the schools full time splitting his time to spend part of it on running the testing programs. Teachers know and like him. There is consensus across all levels, however, that "we're a high pressure district for adults."
- management style characterized by clear reporting channels, teamwork and attention to detail. The Personnel Director, a member of the Administrative Council commented, "It is rare that a single person will make a single decision. It's management by consensus unless you get down to the button." But it is clear that when you get down to the button everyone knows that it is the team of Superintendents who have the final say.

When asked what background experiences or courses influenced their present job performance the most, several central office people cited knowledge of interpersonal relationships. They talked about being aware of other people's time, about "touching base" with one another and about the importance of short turn-around time or decisions.

It is within this context of relative homogeneity of perspective, ten-year staff stability, high level of competence of professionals and managers that the following organizational linkage mechanisms work. These include

committees and regularly scheduled meetings among groups of people as well as connecting relationships among individuals.

Ad Hoc Continua Revision Committees. Groups of teachers and learning specialists get paid during the summer to revise sections of the continua. The Coordinator of materials is responsible for setting up the agenda, providing all the needed materials, and facilitating the task.

The Testing Advisory Committee works with the Testing Director to continually update and improve the CRT system. Over the past year, this Committee has met six times and was composed of one principal and several teachers from different schools. It handled administrative problems, technical problems, log in turn-around time for reporting back. Teachers regularly submit to another group of teachers their comments about items which do not seem to work. The Committee examines the items, revises them as necessary, and reports to the Testing Advisory Committee.

Administrative Council. Weekly meetings of central office personnel to handle problems as they occur. The central office staff meets regularly and formally at weekly Administrative Council meetings. Between meetings they "pop into" one another's office. Some alliances are particularly strong: Between the Superintendent and the Assistant; between the Assistant and the Director of Materials; between the Assistant and the Director of Testing; between the Assistant and the Staff Development Coordinator.

Principals meet together just prior to these meetings to discuss their own problems and concerns.

Communications Council. Monthly meeting with Superintendent, one Board member, one principal, several teacher representatives to share

information and make recommendations. The Superintendents work closely with the Board - in the Communications Council, and individually. They keep them informed of short-term events. They have also worked at defining the relationships. The Board President says, "It takes a long time to learn how to be a Board member. We went to lots of workshops and meetings. We make policy; they run the schools."

Considerable school level interaction occurs. Learning Specialists carry information from one classroom to another as they work with teachers and with groups of children. Learning Specialists conduct schoolwide in-service training. Principals observe in classrooms and discuss lessons with teachers using Professional Development Program language. Three school learning days are allocated during the year.

The Staff Development Coordinator meets twice a month with the Learning Specialists in each school to coordinate staff development activities across the District and to plan new courses based on needs. She also conducts workshops herself. She follows up PDP level one and two courses in the classrooms.

The Materials Coordinator meets monthly with the Media Specialists. She also pays each school a visit early in the school year. She is advised by a teacher committee; and also conducts an annual survey of instructional materials requested by teachers.

Teachers meet with one another informally in their schools, and work formally on the three planning days during one of which they discuss the next year's instructional groupings of their students. They meet with other teachers in the District at professional development courses, at continua-writing or item-writing committees.

The Testing Director is in the schools in his counseling capacity. He works closely with the Learning Specialists to continually debug the CRT operation.

Such networking --in committees, in meetings, and by individuals staying in touch with one another-- seems to be facilitated by:

- smallness of District. The number of people involved are few, the geographic distances between schools and central office are small. Personal contact and communication is a relatively easy matter.
- financial support for these interactions. Teacher released time for up to 7 days a year is a budgeted item. So are the planning days. So is provision for media specialists and new materials.
- common language in which to talk about instruction and instructionally-related services. This is especially important in facilitating principal/learning specialist/teacher interactions; and very beneficial in facilitating testing director/staff development coordinator/materials coordinator interactions.
- shared orientation that the classroom is "where the action is." Central office staff regard themselves as having a service function. They are frequently in the schools and in the classrooms. Conversely, teachers are frequently out of their classrooms.

IMPACT ON CLASSROOM PRACTICES

Classroom practices can be affected by District-level decisions, by school-level decisions, and by individual teachers making classroom-level decisions. In Shelter Grove School District, we have seen a high degree of linkage - in conceptualization and in organizational mechanisms - in the interests of a technically-oriented instructional program. At least two sets of critical questions remain to be asked about what happens in the classroom. The first question asks: HAS CLASSROOM PRACTICE BEEN AFFECTED BY DISTRICT-INTENDED AND OPERATED LINKAGES AND IF SO HOW? It is often the case that connections --however well thought out and however well operationalized at the District level-- are severed when the classroom door closes. Daily life in the classrooms must provide the evidence as to whether the technically-oriented instructional program of Shelter Grove District exists in rhetoric or in actuality.

The second set of questions ask about teachers' overall reactions to District intentions and operations. WHAT ARE THEIR REACTIONS TO TO SPECIFIC FUNCTIONS AND COORDINATING MECHANISMS? WHAT ARE THEIR PERCEPTIONS OF CLASSROOM CONSEQUENCES?

The following analysis is based on interviews with fourteen teachers in three schools, three principals and three learning specialists.

HAS CLASSROOM PRACTICE BEEN AFFECTED BY DISTRICT-INTENDED AND OPERATED LINKAGES BETWEEN TESTING AND INSTRUCTION AND, IF SO, HOW? Without equivocation, for Shelter Grove teachers, the answer is yes.

- Without exception, all the teachers interviewed were aware of the District continuum and the CRT system. They explained in consistent terms the roles of the learning specialist, the

the media specialist, the principal, and the Professional Development Program.

- For most teachers, the continuum and the related CRT direct their selection of content in the basic skills. This is more true for reading and language arts that in math where the textbook sequence is often followed. Sample quotations: "The continuum is a real working tool." "I feel comfortable about using it [the continuum]. My teaching is aimed at it. I also do social action and emotional growth things." "I use CRT in planning. I make a list of areas to work in." "I teach to the test and that's OK. For many, it provides a well thought thruway to organize their teaching. Others like the emphasis on skills." "We teach skills here. How you do it is your business."

One or two teachers reacted against such control or felt that they could not do it all. A teacher in relation to math, said, "I don't let the test influence what I do. I think the continuum has introduced too much in the early grades." A new teacher says, "The first year I just waded through."

- For most teachers, the CRT scores are useful in grouping children and in diagnosing their progress in learning. "The CRTs don't provide too many surprises." "If I've taught it well, kids pass." The teachers welcome the diagnostic screening given to new students by the learning specialist. It helps place them in groups soon after the start of school. Teachers report that student instructional groups change frequently based on CRT results. Often they change within classrooms; sometimes between classes.

The learning specialist facilitates this process by conferencing with teachers after each CRT administration. The media specialists report advising the learning specialists and the teachers on specific student-appropriate material.

- The Professional Development Program gets high marks from teachers. They report that the level one and two courses (in objectives, diagnostic and prescriptive teaching) are not duplicates of what they had in pre-service courses. "PDP makes me aware of what I do. I never got this in college." "Because of the PDP I really make sure to do high-level thinking with the kids - analysis and synthesis." Specially praised were the recent offerings in brain hemisphericity, in written language, in creativity.

WHAT ARE TEACHERS' GENERAL AND SPECIFIC REACTIONS TO THE TEACHING-TESTING-TEACHING INSTRUCTIONAL PROGRAM? WHAT IS THEIR VIEW OF CLASSROOM CONSEQUENCES?

The majority of teachers interviewed have been in the District more than ten years. They indicate that they were first resistant, then skeptical of CRT-continua-diagnostic/prescriptive teaching. Most now are positive. They appear to be positive for a variety of reasons.

1) Pride: "Kids get a phenomenal education here." "This District gets better and better. It keeps changing." "This is a really current District."

2) Professionalism: "There is a push for teachers to keep up on things." "I could take my skills now, and apply them anywhere." "I like the consistency in this District. They really want the kids to learn." "I think teachers should always be growing." "All the teachers are familiar with the same jargon."

3) Feeling of support: "If we have a problem, it's taken care of." "I really like it here. I don't feel the pressure. I've never known anything different." "Administrators want to be helpful." "They know what's going on."

A number of them voice concern about the long-term effects of Shelter Grove's current view of education. These are arranged here from the specific to the general. The ordering does not reflect the frequency of responses.

1) Annoyance with wording of specific items that make results misleading.

2) Amount of time consumed by the tests themselves and by making of the printouts. Respondents said this problem was diminishing as the clerical procedures were being worked out more efficiently.

3) Effects of tests on students. Some children don't like to take tests. Some get very anxious. Some don't take them seriously enough.

4) Tests miss some more important aspects of education - for example, retention and attitudes.

5) The technical approach to teaching is very structured. It may not take into account the learning style of the child, or the teaching style of the adults. (This conversation occurred during discussion of the brain hemisphericity PDP course.)

6) The child gets to make very few choices about his/her own learning under this type of system.

7) There is increasing homogeneity among teachers in the District. Respondents noted that this was good unless carried to extremes.

8) There is continuing pressure on teachers to get better and better. As above, respondents noted that this was good unless carried to extremes. For some, the "burnout" problem seemed very real.

RECENT HISTORY OF SHELTER GROVE

Many of the central office staff, principals, teachers and parents in the Shelter Grove School District have a sense of history. They can remember back to the late '50s when their town was unincorporated, somewhat sleepy and very conservative. Some people remember the town-splitting controversy that erupted when a teacher was accused of being a Communist - this three years after the death of McCarthy. Others remember drug busts in the high school during the early '60s. During those days, though, "the kids were smart" --group IQ tests showed the district's average to be over 110-- and almost all were college-bound.

In slow stages, the community and the school district changed. In 1958, expensive apartment developments brought more people - some professional like the existing residents, others technically-oriented, upwardly mobile engineers or managers who worked in nearby cities. They transformed Shelter Grove into a bedroom community. Throngs of children --whose parents urgently wanted them to learn and to be successful-- crowded into the classrooms. New teachers were hired. Double sessions started. By the mid 1960s, the schools in Shelter Grove seemed to be functioning in the way they always had: large-group instruction in traditional self-contained classrooms, with every teacher on his/her own in interpreting a general subject matter curriculum. But quantity was threatening quality. Conserving meant changing.

In 1965, a new superintendent was hired - aggressive, innovative, interested in expanding the horizons of teachers as well as children. He was controversial. He tried to introduce into this small-town district new ideas in education such as individualized instruction and team teaching.

He put the spotlight on teaching. He tried to reorganize classroom instruction. After four years, he left. But the "system" was not the same afterwards. "He broke us loose," some now say with admiration. "We became aware that teachers were not able to provide individualized instruction. They did not have the skills," says a long-time principal. "We became aware that we didn't know what children were learning, or what teachers were teaching," says a curriculum expert.

The Board looked inside the district for a replacement Superintendent. In 1969 they appointed the then Director of Personnel, previously a principal in the District, to be the new Superintendent - a man who combined firm convictions about the importance of teaching with good management skills. He was not a flamboyant public figure; rather a good inside administrator. Dr. John Carlson immediately appointed as Assistant Superintendent for Curriculum an individual of twenty years' experience as a teacher and curriculum coordinator. Together, they would provide the direction, create the mechanisms, unleash the energies so that by 1980 --eleven years later-- the District would have a well-functioning and self-renewing instructional program.

Although they were not aware of it during their early trials and concomitant errors, conditions were favorable for their success.

- The elected five-member Board was supportive of the newly-appointed Superintendent and Assistant Superintendent.
- This administrative team was familiar with the problems of the District.
- In turn, they had credibility with, and the confidence of, principals and teachers.
- The District had just been "unfrozen" by the previous Superintendent. What had been tried had not worked; but neither

were the old, entrenched ways of doing things still in place.

- Competent people existed within the District.
- Competent consultants were available from outside the District.
- New educational ideas were being discussed within the academic and the professional communities.
- The size of the District was small: 3000 children in seven elementary schools, with seven principals. Communication and coordination were relatively easy.
- The surrounding community was homogeneous in its politics, socioeconomic status, beliefs about education. District administrators could concentrate on resolving educational problems rather than having their energies consumed by political, social and economic issues.
- Parents were uniformly education-oriented. They wanted their children to succeed in school and go on to college. Children's homes, for the most part, reflected a concern for reading, writing and learning.

All did not go smoothly at first. Dr. Carlson and the Assistant Superintendent, Mrs. Lowell, had inherited an orientation towards individualized instruction and were personally supportive of it. But for them, it was not an end in itself. It was a means towards the end of improved learning for children.

As far as respondents within the District can recall, there was and is now no long-term formalized planning process. Indeed, the Superintendent implies that he did not anticipate the need for planning, even though he is characterized

by his colleagues as someone who "thinks ahead" and "likes to know where he is going." Rather, the Superintendent says, "The concept of individualized instruction is deceptively simple. To implement it fully requires rethinking of every aspect of the school program - classroom organization, teacher training, administrative processes, curriculum and instruction.... The need to reorganize in these areas evolved as the District moved towards the objectives of an individualized program." The Assistant Superintendent notes, "We started by wanting to find out what the children were learning. The principals were supposed to be making classroom observations, but they weren't telling us what we wanted to know. So John and I started visiting the classrooms. We found that the teachers couldn't tell whether the children were learning either."

So they looked around for something that would. They found Pacific Learning Corporation in Palo Alto who promised they would write items for district-selected objectives and would provide computer printouts for scores.

Over the summer and into the next school year, a small group of administrators and teachers and outside consultants wrote district objectives. Pacific Learning Corporation generated some items during the following school year. Simultaneously, several teachers "discovered" Madeline Hunter's theories about teaching. The Superintendent encouraged them to attend workshops at the County office.

The first four years of the Carlson-Lowell administration, then, was characterized by exploration. Consultants were brought in, workshops were attended, meetings were held.

In 1970, behavioral objectives were purchased and tried for two years. They turned out to be improperly sequenced. Teachers hated them. They were discarded.

The exploration was stimulated by the ideas-in-mind of the Superintendent and his Assistant. They wanted to improve learning. They had inherited the notion of individualized instruction. The attempt to individualize instruction had unearthed major difficulties both in monitoring children's progress and in managing classroom instruction. One thing led to another. Some things were rejected without trial. Others were rejected after a short trial period. The decision about what to attempt and try to make work were probably made in accordance with the administrators' preferences and their sense of what would work. The Superintendent says now, probably reflecting his similar earlier beliefs, "testing cannot be separated from instruction." The Assistant Superintendent says now, probably reflecting her earlier beliefs, "staff development is at the heart of the instructional program."

In 1973, the first Pacific Learning Corporation criterion-referenced tests, known as PAL (Pupil Assessment Lab) were administered. They were intended to monitor student learning so as to feed back information to teachers about how individualized instruction was working. They were also seen as a reassurance to anxious parents who were worrying that their children were not acquiring the basic skills.

There is general agreement among district respondents with the teacher who said "it was a mess." Everyone complained. Teachers clearly remembered their complaints:

- it was a waste of time. It didn't tell us anything we needed to know.
- The children were confused by the tests.
- The first two weeks of school was a bad time to be testing.
- We didn't know what to do with the results.
- We didn't want to be judged on the basis of what our children did on those tests.

• It was not coordinated with anything we taught.

Teachers made their complaints known to parents and to the Board. It was a hard time for central office staff. They wanted the CRT system to work.

They appointed teacher committees to try to revise the items. They responded to the need for coordinating testing and curriculum by beginning to work on a district level continuum. A group of volunteer teachers were paid to work on a reading continuum during the summer. Their work continued through the school year.

By 1976, there was a reading continuum (RIC) and the beginning of a math continuum (MIC), both teacher-committee generated. By 1976, the PAL criterion-referenced testing system had been scrapped. Teacher committees had generated reading, math and language arts test items to form their own Criterion-Referenced Instructional Tests (CRIT).

Although teachers remember those early years of the '70s as a time of intense frustration, they generally agree with the learning specialist who said, "If it hadn't been for PAL, we wouldn't have moved to CRIT and the continua. We initially thought we were making a continuum to fit PAL. We made the continuum, and then scrapped PAL." And with another learning specialist who, after describing the process of continuum development and item writing, said "Most people welcomed the chance to get into something."

By 1976, too, the explorations into staff development activities to improve teaching had evolved into the currently-in-place Professional Development Program. Building on the interest of a few teachers in the Madeline Hunter workshops, and the intense conviction of Mrs. Lowell that teachers needed a way of thinking about teaching in order to make individualized instruction work, the District had hired a Professional Development

Coordinator. Starting with interested teachers, she developed Hunter-type workshops. At first, the program was voluntary. Encouraged by teacher interest and by the coincidence of viewpoint with the state-adopted Early Childhood Program, the District soon made the PDP mandatory for all principals, and for all new teachers. It soon became the "expected" thing for all teachers to participate in. Because teachers were comfortable about leaving their classes to attend PDP courses, the training was gradually extended to include substitutes and aides.*

Other elements of the support structure needed to link testing with instruction in the interests of teaching every child evolved in a similar one-step-after-the-other fashion. In addition to the continua, the CRT system and the PDP, those elements include a Learning Specialist in each school, media and materials facilities, clinical supervision.*

From its recent history, and given a continuation of the same conditions, we would expect Shelter Grove to continue evolving. Some of the directions have been pointed out by respondents in response to questions about problems in the current operations of the District. They include:

- attention to the potential narrowing of the curriculum to only those subjects in the continua and in the CRT program
- attention to the possibility of reducing educational aims to those which can be compartmentalized into objectives and measured by tests of some sort
- need to coordinate the testing system with other requirements coming from outside, such as proficiency tests, state assessment tests
- need to fine-tune test items, reexamine item clusters and develop criteria for setting standards

---

* Details about these presently-functioning elements are found in an earlier section.

Conditions, however, are not likely to remain during the 1980s as they were during the 1970s. Among the large-scale changes likely to influence District direction are:

- the departure of both the Superintendent and the Assistant Superintendent.
- the expansion of the District from an elementary to a unified district to include two middle schools and a high school.
- the potential loss of federal and state fundings due to changes in laws and regulations.

A review of the past fifteen years of District history cannot predict the future. But there are some notable characteristics about the evolution of a well-developed, continually renewing that are worth noting. While retrospectively, one could impose a rational order to the way in which events occurred in Shelter Grove, it is the recall of the respondents who lived through this period that there was no master plan. There was instead, a general direction in which the administrators wanted to go. There seemed not to be a conscious weighing of all possibilities and the subsequent selection of "the best." Rather, there were two levels at which actions, on an individual basis, were tried out. The first level could be called an "imaged" trial, where the errors became evident as soon as the action was mentally conceived. The second level was an actual tryout where the events following the action took on a life and influence of their own. In some cases, the errors led to abandoning the effort as in the District's purchase of behavioral objectives. In other cases, the errors led to modifications such as in the revision of the early PAL tests and subsequent ramification into the current continua and CRIT.

The trials and errors were not without costs - of frustration, of discontent, and of time and money. Nonetheless, these "hard times" were

somehow transformed into a shared ordeal in the interests of a common purpose. Whether this is the gloss of memory, or whether the difficulties actually built a team spirit at the time is hard to determine. What is evident now is that, within the District, there is a high level of self-confidence and of professional power. It is possible that morale was built from the energies released in mutually coping with the problems of making instruction better, or, that morale was the sustaining force that helped people endure through changes. Whichever the explanation, some of the trials resulted in successes and were incorporated in workable and integrated fashion into the District's operation.

What is the significance of particular people in bringing this ten-year evolutionary process of linking testing with instruction to fruition? According to all our District respondents, the key individuals were the Superintendent and the Assistant Superintendent who were the original owners of the vision - and who created the circumstances under which the Board, the remainder of the central office staff, the principals and the teachers shared in the vision. They appeared to have possessed the managerial and interpersonal skills to move the process along - not without a cost, a casualty, and not without mistakes and wrong turns - but in a generally consistent direction.

STILTON UNIFIED SCHOOL DISTRICT
1980

BACKGROUND INFORMATION

The Stilton Unified School District is a medium-sized district in a community undergoing rapid transition and population growth. From a primarily blue-collar and semirural community five years ago, the town of Stilton is changing into white-collar and professional. From a single industry town Stilton is becoming a bedroom community to the large metropolitan area seventy-five miles away.

The Stilton Unified School District contains thirteen elementary schools, three junior high schools serving grades 7, 8 and 9, one high school serving grades 10, 11 and 12, and one continuation high school. There are 12,000 students attending these schools at the last count, although the population may have increased subsequently. There are 623 certificated personnel and an additional 211 aides.

Stilton is a "low wealth" school district. Like other such districts within the state, however, it receives approximately two million dollars annually from the state for School Improvement and State Compensatory Education programs. The district also receives federal funds through the Title I program.

Within the last few years the Stilton School District has found what seems to it the solution to problems of low student achievement, community dissatisfaction with the school system, and rising aspirations on the part of the heterogeneous population within the community. Stilton has embarked upon the transformation of its elementary schools into fundamental schools. Although the term fundamental school means different things to different people, and although the supporting systems for what may be thought of as a massive change in direction are not yet in place, the community and many individuals in the district are enthusiastic and excited by the prospect.

RECENT HISTORY OF STILTON

During the Second World War Stilton was an agricultural region. There were few houses, some orange groves, grape vineyards, and chicken farms. After the war, a major steel company built a production plant in Stilton. The community almost overnight changed from rural to industrial. The community, however, was not a stable one. Not only was it made up of new arrivals; the new arrivals themselves were extremely transient. As workers within the steel industry they were transferred between mills within the company's system.

During the middle and late sixties Stilton received a great deal of government money especially in housing and welfare as a result of President Johnson's push for a Great Society. The economy of Stilton flourished and the population increased.

Another drastic shift in the nature of Stilton's population occurred in the mid-seventies. White-collar workers and professionals from the metropolitan area moved into Stilton transforming half of it into a bedroom community. The new arrivals are wealthier, upwardly mobile, educationally-oriented. In the past two years the socioeconomic status of Stilton has increased sharply. The percentage of students receiving Aid for Dependent Children has declined.

The School District has changed in accordance with the changes in the community. During the fifties the School District emphasized vocational skills, but each teacher "did his or her own thing." The classroom doors were effectively closed to the central office administrators. The central office itself was very small. They provided assistance to schools, for example, with audiovisual materials and general administrative support.

In 1956 the District became unified, incorporating the junior high and high schools into its aegis. During the 1960's, Stilton professionals reacted to the innovative forces within American education. There was some experimentation with team teaching, with individualized instruction, with multiage grouping. The District Office provided neither a particular direction for this experimentation nor particular support for it. Every school operated with a good deal of autonomy. Between 1967 and 1972 there were five different Superintendents. Each of these Superintendents wanted to leave his mark on the District and emphasized a different aspect of curricular or instructional change. Each Superintendent was implementing policies more in response to his own educational philosophy and professional view rather than in accordance with what the community wanted. The Board of Education translated the community's lack of faith or lack of understanding of these educational ideas into a revolving door for Superintendents.

The current Superintendent, appointed in 1972, began his tenure with an emphasis on individualized instruction. However, because of discontent on the part of the community and the Board with low test scores and with other evidences of students' learning deficiencies, individualization has given way to an emphasis on basic skills organized according to grade-level standards. Accompanying the emphasis on basic skills has been a commitment to certain traditional features of schooling.

EDUCATIONAL FOCUS

Stilton Unified School District is moving from an individualized approach to education, matching the instructional program to student needs, to a fundamental, back-to-basics approach in its thirteen elementary schools. This effort is tacitly supported by the Superintendent, and actively supported by the Board. This direction is characterized by an emphasis on

- .student attainment of basic skills in grammar, reading, arithmetic by all students
- .whole group instruction in contrast with small group or individualized instruction
- .elimination of combination classrooms in favor of grade rooms
- .homework and sanctions for not completing it
- .the retention of students who do not pass grade-level requirements
- .discipline with the reinstitution of corporal punishment if authorized by parents, or suspension if not
- .dress code
- .attention to matters such as deportment, attendance, etc.

During the 1978-79 school year, one elementary school instituted the fundamental approach. Teachers and students from all over the District participated on a voluntary basis. Those previously at the school who disagreed with the philosophy of the experiment were relocated elsewhere in the district. At the end of the year the Board of Education proclaimed enthusiastically that the experiment was a success. Student test scores had risen, although the extent to which such rise was due to fundamental-type school activities rather than to alternative

interpretations such as changes in population, or to the heady feelings accompanying participation in an experiment, was not examined. The Board of Education was pleased. Parents of enrolled students were delighted. School attendance had improved and vandalism had declined. The school's principal gained much popular support.

During the 1979-80 school year, the District authorized that five additional schools be brought into the program. Results of that year's efforts were more mixed since neither students nor teachers at the five schools were voluntary. Our respondents told us that attitudes more than activities had changed. Student test scores have not yet been analyzed. Nonetheless, many of the Central Office staff believed that they were on the right path to educational improvement. Strongly supported by the Board, all of Stilton's elementary schools will adopt the fundamental school emphases outlined earlier for the 1980-81 school year.

Testing activities within Stilton predated the current orientation towards fundamental schools. The Comprehensive Tests of Basic Skills (CTBS) has been used for many years in order to fulfill federal evaluation requirements. Beginning about six years ago, the district has also administered the Survey of Basic Skills at grades 3, 6 and 12; those tests are part of the Assessment Program mandated by the State Department of Education. The district is now in the process of developing the tests and the remedial programs needed for the state-mandated minimum competencies testing. Currently, also, Stilton schools which participate in the state-funded School Improvement Program are visited by Program Quality Review teams who asses the extent of school-site planning and the consistency of

activities with previously developed plans. However, at present, the ideas surrounding the use of test specifications and of test scores in Stilton are linked to the fundamental school orientation.

- test specifications, specifically the CTBS, may become the organizers of the basic skills content and skills to be covered in each grade level

- test scores, specifically the state Assessment Program tests, may become the indicators of the success of the fundamental school emphasis

- scores, both CTBS and State Assessment, may become methods for identifying those schools, grade levels or teachers associated with deficiencies in student learning.

Evaluation, as distinct from testing, also seems to be a salient concern in Stilton. Although not explicitly stated by respondents, evaluation may become the vehicle for monitoring school compliance with central office emphasis on fundamental education. The District in the Spring of 1980 organized its own review teams to visit schools and discuss with principals and staff aspects of school programming.

The following sections describe some of the details of the District's anticipated functioning in the areas of testing, evaluation and instruction. Coordinating mechanisms will be described as well as evidence of impact on teachers.

It must be remembered that Stilton is a District in transition. The unifying fundamental school idea is only two years old and is just now being implemented in all District elementary schools. There was considerable anxiety and concern expressed by some respondents about the advisability of this approach. However, the strong support by the Board and

and the community of the new direction makes it likely that Stilton will proceed to coordinate District functions so they support the complex values repesented by the terms "fundamental schools" and "back to basics."

DESCRIPTION OF DISTRICT FUNCTIONS

Testing and evaluation. Successful instruction in Stilton is now defined as that which raises test scores. This appears to be the position of many people in the central office, although there are variations in their levels of commitment. Those who are most knowledgeable about the details of test construction and the ambiguities of test analysis raise the most questions about this position. Principals and teachers, as a group, are also less enthusiastic than central office staff about organizing instruction around "teaching to the test."

The tests which are to act as the primary indicators of student learning are the CTBS and the State Assessment Program tests. Since these tests emphasize reading, math and language arts, the curriculum is focused on these three subject areas with diminution of attention to social studies, science and the arts.

The union of test scores with fundamental schools represents an uneasy alliance of the ideas of two influential men within the District: the former principal who seemed to have demonstrated in his own school that the solution to the educational problems of the District was in a fundamental no-nonsense approach to learning; and the Assistant Superintendent, who was attempting to develop a systematic teaching-testing-teaching procedure for individualizing instruction before the principal's solution seized people's imagination.

The systematic teaching-testing-teaching procedure was to be keyed to criterion-referenced tests rather than to the CTBS or to the State Assessment tests. The Assistant Superintendent had initiated the use of McGraw-Hill's Individualized Criterion-Referenced Testing (ICRIT) system

for reading on a Districtwide basis. He had also urged each school to develop its own criterion-referenced tests in math and language arts. The basis for these was to be the continuum in math and language arts developed K-8 by teachers acting under the direction of one of his trusted associates.

New ideas about fundamental schools postponed the development of the District continua and the CRT system for approximately one year. A new development committee, under the direction of a fundamental school support, revised the continua and is now in the process of integrating the individual school CRT tests into a Districtwide testing system. New District-developed reading tests are replacing the unpopular McGraw-Hill ICRIT. The Districtwide CRTs when completed will consist of five questions for each of the forty skills for each subject for each grade.

It is the responsibility of the Testing Coordinator in Stilton to review test results obtained from the CTBS and from the State Assessment Program tests school by school. Once each year, he meets with principals and teachers to review the implications of the scores for school-site planning. We do not yet have information about follow-up or impact of these school visits on teacher activities in the classroom.

The District's interest in the use of testing and evaluation to structure curriculum and to monitor school-site functioning is further illustrated by the District's evaluation review teams, first started in Spring 1980, which are to visit the schools once a year. One respondent described the wide-ranging interests of these teams as including

- the learning atmoshpere for students
- the feelings of students
- the services provided by aides

- the communication between teachers and aides
- the materials used in the classroom
- the classroom management skills of the teacher

The review team conducts an exit interview with the principal and staff. Follow-up appears to be in the hands of the principal, with monitoring of their actions left to informal interaction between the Test Coordinator and the individual principal.

Summary. In Stilton, the District intention appears to be to link testing and evaluation closely with instruction. Test specifications are being used to rethink the curriculum. Test scores are being used to monitor student performance, grade-level performances, school performance. Teachers express concern that the scores will be used to assess teacher performance as well although District officials say that will not happen. The measure of the success of education in Stilton is portrayed in terms of student performance on test scores. Evaluation teams visit the schools to assess and discuss school functioning with principals and staff.

From our interviews, we do not know whether Stilton has yet thought about how staff development and materials selection might support their testing/evaluation/instruction nexus. Staff development is currently handled at the school level with the exception of a single Districtwide in-service day. Materials selection is likewise handled at the school level.

DESCRIPTION OF LINKING MECHANISMS

Presently, Stilton is experiencing an ideological schism in its approach to education. In one camp are the proponents who advocate a fundamental, back-to-basics approach utilizing large group instruction and defining specific skills to be mastered at each grade level. Students failing to master these signpost skills will be retained at their grade level. This approach is heavily behavioristic, that is, reward and punishment-oriented. Teachers are to spend a lot of time on only three primary subjects in the elementary schools: reading, language and math.

In the other camp, which at this particular moment in time is in a subordinate role to the proponents of the fundamental methodology, are strong advocates of a cognitive approach to instruction. The proponents of this camp favor an individualized approach to education, matching the instructional program to student needs. These people had originally developed the District continuum for grades K-8 in reading, language and math. That has now been reformulated by the behaviorists. The behaviorists, utilizing the reformulated continuum, have developed a management system consisting of the continuum, CRTs to measure each signpost skill, and a computer support system. From this management system, the behaviorists intend to measure the progress of student achievement and to be assured that they have mastered all the skills designated at their grade level before promotion. Presently, the resolution between the behaviorists and cognitives is not at hand, but it appears that both are utilizing the management system as a conciliatory apparatus to resolve their differences.

In addition to the management system --which is not yet fully operational-- the fundamentalists see several other coordinating mechanisms. They include a strong role for the Director of Testing and Evaluation who can link District policies with schools' activities. They also inculde a stronger role for the Program Manager's biweekly meetings. These individuals are the Title I coordinators in each school who also perform assistant principal duties. Communication among them would link schools together in moving forward on the fundamentalist road. A 1978 audit recommended that this group investigate the possibilities for Districtwide in-services, or, if that proved not feasible, sharing of in-service content among a subset of schools.

Other formal mechanisms are likely. For the present, however, the personality and powers of the popularly support principal seem to be persuasive. Applying the same techniques of clearly defined rules accompanied by standardized rewards and punishments that he wants teachers to use with students to the teachers and principals themselves, he has created an atmosphere where opposing ideas are not openly expressed or explored.

IMPACT ON CLASSROOM PRACTICE

Only four teachers, two principals and one program manager were interviewed about the impact of District testing and evaluation on their instructional activities. The four teachers were aware of the recent changes in District orientation. They pointed out that "expectations of teachers and students are now much higher, "and "I'm a little more structured now than before." Also, "I retain children in grade now more than I used to." Three of the four had participated in in-service courses dealing with assertive discipline. The three administrators mentioned the new emphasis on "academic subjects," "basic skills," and on "directed lessons." The four teachers spoke approvingly of the District continuum indicating that it provided some "order" in their thinking about teaching reading and math skills. The management system requiring CRT testing of benchmark objectives, probably because of its newness, was rated as having only a little or some influence.

When asked about actual changes in their classroom practices now as compared with several years ago, two of the teachers mentioned differences which they attributed to their own experience or to changes in textbooks. None of the respondents saw the District's intentions as controlling of their classroom behaviors, nor did they see their principals as closely supervising their instructional strategies. This was probably an expected result since the fundamental orientation had not, by the time of the interviews, spread to all schools.

Teachers differed in their estimate of the influence of the CTBS test on their thinking or behaviors; the three administrators were aware of its use as an identifier of student deficiencies and of its tendency to propel teachers "to teach to the test."

Summary. It is not possible to describe the classroom impact of testing and evaluation from short interviews with seven administrators and teachers at the end of the school year 1980. As noted earlier, the relationships among testing and evaluation and instruction are mediated through the ideology of the fundamental school. Fundamental schools have been tried in one school for two years and five schools for one year; they will affect all schools for the first time in 1980-81. Only after some time has passed will classroom changes become assessable.

APPLICATIONS OF ANALYTIC FRAMEWORK

BAYVIEW

Summary

Bayview central office administrators and the teachers in Bayview's twenty-two schools are in general agreement that neither mandated tests nor the evaluation requirements attached to categorical programs should drive their efforts at instructional improvement. Rather, those we interviewed emphasized their belief that more effective teaching and therefore increased student learning would result from encouraging teachers to use techniques such as task analysis and diagnostic prescriptive strategies. District office personnel stress, in addition, the importance of the principal's role in providing clinical supervision for teachers and in creating a school climate that fosters professional teamwork.

Bayview's six-year thrust towards staff development has been propelled by the interest of particular individuals within the district, most of whom had shared experiences in a single elementary school. Additionally, the Superintendent who served from 1972 to 1980 encouraged the writing of grants to procure federal and state funds for staff development activities.

Recently two factors have stimulated the district staff to think about integrating tests with instruction. The first is their experience with state-mandated, district development proficiency tests which students must pass in order to receive a high school diploma. The anticipation of parent protests in response to the denial of diplomas has prompted the creation of remedial programs, new staff development activities and the development of districtwide continua for reading, language arts and math. The second factor is the availability of technical assistance from the state office overseeing State Assessment Program testing. Such technical assistance has provided the direction for the district to analyze their curriculum in relation to SAP test specifications.

The development of a T/E/I linking subsystem is just beginning in Bayview. Although many of the operations and coordinating mechanisms are in place, a coherent set of ideas to guide the development of the system is just starting to emerge.

Discussion Using Analytic Framework Categories

The two environments most relevant to Bayview's development of a T/E/I linking subsystem are the local community and federal and state requirements. The diversity of the school-age population in terms of their socioeconomic status, their language proficiency and their

educational aspirations contributed to the district's decentralized management of curriculum and instruction. It has also made teacher capacity building through staff development a reasonable vehicle for instructional improvement.

On the other hand, in recent years, the state has required the administration of State Assessment Program tests and district-developed proficiency examinations. Schoolwide scores on the State Assessment Program tests are released to the press concurrently with their transmission to the district. The School Board has taken keen interest in these scores. The districtwide proficiency examination, starting in June 1981 must be passed by students in order to receive a diploma. The anticipation of parent reactions has also encouraged the district to take action. The linkage of these two tests with instruction is just beginning in Bayview.

There is, among Bayview district personnel, a coherent set of ideas relating staff development for both teachers and principals to instructional improvement. This set of ideas forms the conceptual core of Bayview's efforts. Ideas about how to use tests, to analyze curriculum, to review materials, to analyze teacher strategies, are fragmented.

Nonetheless, new activities are under way. District continua in math, reading and language arts are being developed by teacher committees. A criterion-referenced testing system is being constructed to match the continua. CTBS testing is being resumed after almost a decade's hiatus.

Concurrent with these activities, new roles and relationships within the central office are evolving. A new Coordinator for Testing

is working more closely with the individuals responsible for curriculum and staff development. The evaluator of specially-funded programs is attempting to create a climate where his evaluation results will receive increased attention.

Within the central office, there is a sense of commitment to instructional improvement through staff development. This commitment appears to be shared by teachers in Bayview's schools although the extent to which classroom activities have changed as a result of staff development is unclear. Bayview's T/E/I linking subsystem is in its early stages. Its growth depends very much on how the leadership component within the central office assess the benefits likely to accrue from such a system.

NORTHTOWN

Northtown T/E/I linking subsystems evolved out of the district's previous efforts in the 1960's to develop a decentralized organizational structure that maximized local school site determination of the instructional program. Working in a relatively benign community environment, the district recruited very capable staff, developed an elaborate cross-functional committee structure, and established a comprehensive in-service training program. The evaluation office gained district-wide credability with its work on establishing a link between the time on task research and school effectiveness.

Northtown has begun linking CTBS test scores with its decentralized school site decision-making system through the use of annual evaluation reports to each school. The Staff's desire that tests other than CTBS might be considered for use in improving instruction, e.g., CRTs, have had to be abandoned because the courts and some community groups have identified CTBS results as the measure of the district's competency and commitment. Further, under court pressure, the district has had to modify its decentralized view of instructional decision-making in selected racially isolated minority schools and implement centrally determined instructional programs with predetermined standards of pupil progress as measured by CTBS gains. Northtown is now struggling to maintain simultaneously two instructional decision-making structures The development of the T/E/I subsystem is in a holding pattern or may evolve in some new, yet undetermined direction.

Discussion Using Analytic Framework Categories

While Northtown seems to have made considerable progress in assembling the necessary components for an effective T/E/I linking subsystem much remains to be done.

In Northtown we have a clear illustration of how the interaction between the district and its environment influenced the T/E/I linking subsystem. The components of that subsystem can be traced to the early 1960's when the district population was growing rapidly and administrators formed a management task force to help plan for the future. At that time the community --a relevant environment-- was either supportive of or neutral toward the district's program. Neither the Board, nor groups

of parents crossed organizational boundaries to advocate specific remedies or directions. In this relatively benign environment, the district was free to adopt their consultant firm's recommendation to develop a decentralized management plan that promoted considerable local school site responsibility for instructional programs. In recognition of the proliferation of instructional programs that might occur, an extensive integrative mechanism in the form of a network of cross-functional committees was developed. This structure, designed to maximize the involvement of many constituencies in districtwide decisionmaking, served as a communication device and helped build staff commitment for the district's programs.

In addition the district established an extensive in-service training program that supported school site efforts to develop and continually improve the appropriate school site instructional program.

This decentralized approach was strengthened by the district's practice of recruiting nationwide for professional staff of exceptionally high caliber. The district's unusually capable and cosmopolitan staff is aware of University-based thinking and contributes to national professional and research activities.

The evaluation office forms the nucleus of the district's T/E/I linking subsystem. It was the district's operation of statewide programs in early childhood education. This activity gave the evaluation office considerable district visibility and enhanced its credibility with teachers. The current role of the evaluation office staff in district decisionmaking is very important.

Thus, many of the elements of Northtown's T/E/I linking subsystem have flowed from the original idea of building and managing a decentralized school district.

Recently, external events have had a profound effect on the district's evaluation and testing programs. Required to use the norm-referenced CTBS tests to satisfy certain judicial mandates, the district has built a T/E/I linking subsystem which utilizes these tests. This subsystem, though not operating in all schools, is an attempt to link student scores on norm-referenced tests to local school-site instructional decisionmaking through the mechanism of evaluation reports. The preexisting mechanisms have allowed the district to begin building staff understanding of and agreement with this program. The in-service component supports the program and a competent staff is available to implement it. It should be noted that all these parts have fallen into place after the fact. The T/E/I linking subsystem has essentially emerged from other programs and priorities.

The system is by no means fully operating or complete. The evaluation office, for example, has virtually no role in the development and assessment of the district's CRTs and the proficiency testing program. Although there is interest in more closely linking the district's instructional continuum to a criterion-referenced testing program, the court, and significant segments of the public, have now fixated on CTBS scores as <u>the</u> measure of the district's competency and commitment toward solving the problems of racial isolation. Given the possibility of forced busing if the district does not respond to the court's mandates, the district's energies are being directed at meeting externally-imposed standards, standards which some district staff feel are unproductive

and unrealistic. Thus the district's T/E/I subsystem development is in either a holding pattern or about to embark on a new, unexpected direction. The T/E/I linking subsystem's continued evolution and change will be likely the result of a combination of district planning and unexpected and uncontrollable external influences.

SHELTER GROVE

Summary

Shelter Grove's T/E/I linking subsystem is aimed at keeping track of the learning of each child in the district. Almost all of the district's operations are thought of as a support system for teachers' use of a teaching-testing-teaching cycle. Those individuals at each level --classroom, school, district, community-- appear to share the view that this arrangement produces a high level of learning for students.

Three factors seem to have facilitated the emergence of Shelter Grove's T/E/I subsystem: the support of a homogeneous education-oriented community; the stability and cohesiveness of the teaching and the administrative staff; and the relative lack of social or political problems besetting the district.

The motivation for the emergence of the T/E/I subsystem seems to have been the ideas and vision of the Superintendent and the Associate Superintendent augmented by the availability of conceptual and technical resources in the form of competent individuals within the district and outside consultants, products, or systems that the district could draw from. (Interestingly, the availability of government funding or the requirements of federal or state programs do not seem to have played an important part in the development or maintenance of Shelter Grove's T/E/I subsystem.)

The eight-year effort to develop the T/E/I subsystem was evolutionary; it can be more accurately described as trial and error --with one thing leading to another-- than a phased programmatic effort directed by a long-range plan. The subsystem, according to district respondents, is "in place" now. The major task confronting a newly expanded unified school district in the future is the extension of the teaching-testing-teaching cycle to secondary schools.

Discussion Using Analytic Framework Categories

Environments relevant for district administrators in developing a T/E/I linking subsystem in Shelter Grove were first, their immediate community and secondly, their professional connections. Shelter Grove, for the past eight years, has been a small elementary district consisting of seven schools and principals, one-hundred thirty-two teachers and three thousand children. The school system provided a sense of

community for many parents --mostly college-educated and in technical or professional positions-- who lived in Shelter Grove but worked in various locations up to fifty miles distant. These parents met one another at school functions, were interested in a good education for their children, and supported the district in its efforts to achieve a well-managed, individualized instructional system. Not only were the parents united in their view of education and similar to one another in social and political outlook; the entire area remained insulated from many of the social changes impacting other communities. Population increases were gradual, not sudden. The small number of minority children spoke English and were long-term residents of the community. There was no outside intervention from the courts or from state or federal agencies.

Although this protected, benign setting might have led to educational stagnation it instead made possible the gradual introduction of new ideas related to the management of instruction. The Superintendent and Associate Superintendent, appointed in 1972 from within the District Office, began by asking a basic question: "What are our students learning?" They sought methods for answering that question from a commercial publisher of a criterion-referenced testing system as well as from consultants attached to the County Office and to nearby universities. Ideas from within the academic community about measurable objectives and research-based principles of instruction clearly influenced the sequence of events in Shelter Grove.

That sequence included the purchase of a CRT system, the abandonment of that system, the development of district continua in math,

language arts and reading by teacher committees, the subsequent creation of a local CRT system. To support these core operations, staff development courses to enable teachers to teach the skills tested and principals to supervise that teaching, were made mandatory; learning specialists were assigned to each school to work directly with students and to advise teachers on classroom management techniques, and each school had a materials and media specialist to advise teachers on texts and supplements to meet individual needs.

The small number of principals, the small number of administrators, regular management meetings, the informal style of the Superintendent and Associate Superintendent, all contributed to the spread of the central idea of the district --that of linking all operations to the instructional needs of individual learners. The required staff development activities and their participation on development committees socialized teachers into using strategies coordinated with the CRTs tests. Initial teacher resistance was overcome by the persistence of the central office staff and by their responsiveness in dealing with troublesome particulars s ch as inappropriate items, cumbersome paperwork procedures.

While the locally developed continua and criterion-referenced tests are at the heart of the T/E/I linking subsystem, other tests and evaluations required by agencies outside the district are administered by the district. These tests and evaluations differ in the saliency they have for district administrators and teachers. CTBS tests are given to the 200 children in the two schools receiving Title I funding, but only modest attention is paid to student scores. The State

Assessment Program tests are administered to grades 1, 3 and 6. Students generally score high. The Director of Testing identifies subject matter areas where there seems to be a decline in scores from one year to the next. Possible reasons for the decline are proffered and the district considers whether to take any remedial action through staff development activities. The district participates in the state-funded School Improvement Program whose emphasis is on school-site program planning and evaluation. Administrators and respondents say about the program, "We were there before they were." They find the requirements congenial with their view of teaching, and they take seriously the Program Quality Revision teams who evaluate by means of three-day on-site visits. These evaluations appear to shape school-level programs because teachers and principals anticipate the findings of the teams and correct deficiencies in advance of the actual visits.

The impact of the district's T/E/I linkage subsystem on teachers' thinking and classroom behaviors appears considerable. Teachers responded consistently with one another to questions about how they organized their instruction, what influenced their selection of content, how they used test scores, etc.

Whether the link between the CRT system and teaching can be expanded to the secondary schools newly unified with Shelter Grove's seven elementary schools remains to be seen. Of interest also is the extent to which the T/E/I linking subsystem can be maintained without the Superintendent and Associate Superintendent, both of whom left the district in June, 1980.

STILTON

Summary

Stilton's primary emphasis, at the present time, is in converting its thirteen elementary schools from their customary operations into schools with a common fundamental philosophy. "Fundamental" has not yet been translated into operational terms, and it is likely that diversity among the schools, especially at the classroom level, will remain for some time. However, the district hopes to move all schools towards an emphasis on homework, student responsibility, grade-level standards for reading, writing and math.

With this effort, district administrators hope to respond to their Board's wishes that student scores on the norm-referenced CTBS test increase. Consequently, they have initiated this year an internal management system. Teacher committees developed continua in math, reading and language arts. Other groups assembled tests of five items each to measure specific skills listed on the continua. The Director of Testing hopes that these tests can monitor student learning. However, other supportive mechanisms which might encourage teachers to change their teaching strategies or the materials and texts they use are not now in place.

Stilton administrators have not anchored their instructional important efforts to either the CTBS or to the CRT system. They speak of using the State Assessment Program scores to review annually with each school their site-planning effort for the coming year. They also have assembled evaluation review teams to visit schools and report back to the principal their impressions on a wide range of activities.

Testing, evaluation and fundamental schools are all salient topics for Stilton administrators. Although individuals speak of one or several of the elements which might comprise a T/E/I linking subsystem, a coordinated set of operations informed by a conceptualization of how they should operate is not yet in place in Stilton.

Discussion Using Analytic Framework Categories

The Stilton school district is responsive to the educational orientation of its immediate community. This orientation has shifted over the recent past reflecting changes in the larger public's preferences about education as well as changes in the socioeconomic composition of parents. For example, during the '50s, respondents who had lived in the district for many years said that every teacher did his/her own "thing"; during the '60s, there were attempts at team teaching, individualized instruction, multiage grouping. However, Stilton's predominantly blue-collar population wasn't happy with the educational system; the Board of Education translated the community's lack of faith in these educational ideas into a revolving door for Superintendents --five in five years. The present Superintendent came to the District in 1972 intending to implement individualized instruction. He now supports an emphasis on basic skills, whole group instruction, homework, grade-level standards with retention for those who fail.

Stilton's move to fundamental schools and to an increased emphasis on test scores as a measure of success of the district in teaching students to read and write has not been importantly influenced either by research findings or by the views of consultants on education. Stilton

administrators did visit a neighboring district's fundamental schools. They also hired a consulting firm to advise them on job descriptions and ways to allocate roles and responsibilities.

Administrators within Stilton have many ideas about the improvement of instruction and the role of tests and evaluations. However, their sometimes conflicting ideas have not been brought to consensus or coherence. There is, as yet, no attempt to bring district operations such as staff development, materials section, personnel, budgeting into an arrangement supportive of instructional change. It is also too early to say whether the new management plan with its redefinition of functions can coexist effectively with the informal and long-standing relationships among staff.

OBSERVATIONS FROM FOUR CASE STUDY DISTRICTS

OBSERVATIONS FROM FOUR CASE STUDY DISTRICTS

In the following discussion, we first cluster our observations derived from the fieldwork in four districts into categories.

The first three clusters of observations provide a picture of district context, of testing activities within the districts, and of evaluation activities within the districts. The remaining cluster describes the T/E/I linking subsystems in the districts, alludes to the ideas, operations and coordinating mechanisms that comprise elements of the subsystems and assesses the influence of various external environments on the development of those subsystems. Finally we deal with the issue of impact of such T/E/I subsystems on teachers and classroom activities.

Within these categories we present, in discursive fashion and at an intermediate level of abstraction, some of the similarities as well as the differences we observed among districts. Our points of view have emerged from extensive conversations and reconsiderations of both the fieldwork and from a review of the literature. We have included examples throughout. However, details have sometimes been elided for brevity. The reader should refer to the longer case study documents for additional corroboration.

ABOUT DISTRICT CONTEXT

One of the most striking similarities we noticed about our four districts was the amount of turmoil they were experiencing. One district was preoccupied with responding to court directives which will necessitate large management changes in the district as a whole and in the amount of autonomy teachers in minority isolated schools have in instructional decision-making.

Another district was moving to change all of its elementary schools to fundamental schools after a brief two year period of pilot testing the idea in a few schools. A third district was struggling with increases in population which added to the already diverse mix of students. Organizing effective instruction for minority children with limited English-speaking capability was a particular problem in this district. The fourth district, while not experiencing these particular societal tensions, was trying to cope with the incorporation of secondary schools into the existing elementary school district. In addition to trying to deal with these immediate problems, school officials felt no certainty that solutions invented for this year's probelms would be adequate for next year's problems.

In each of these districts one can observe evidence of goal diffuseness and boundary permeability. The external environment often crosses district boundries, e.g., court mandates, demands for bilingual programs, and force the districts to redirect their energies in new or somewhat altered directions. This results in goal diffuseness. District officials cannot maintain consistency over time in what goals to pursue and what priorities to assign. Given this, the district's ability to develop and plan long term is challenged.

In view of these factors, it would be surprising if testing and evaluation were uppermost in the minds of school officials. While not their first concern, however, in these four districts testing and evaluation activities and their linkages to instructional improvement are indeed on the minds of administrators, and have been for some time. This appears due to the convergence of a number of factors: the impact of outside environment on the district, including current federal and state requirements, and the professional biases of key personnel.

In these four districts, then, in spite of the turmoil that they share with other school districts, goal diffusiveness is not total. The linkage of testing with instruction has been an idea in the minds of district personnel for some time.

Another notable characteristic of the four districts we have studied, then, is the professional biases of key personnel towards the improvement of instruction, even though the nature of instructional improvement differs from district to district. In one district a centralized T/E/I linking subsystem evolved slowly over the course of eight years and is now in place. In two other districts, both with larger and more heterogeneous populations, T/E/I linkages are taking place within otherwise decentralized systems. In these districts administrators and teachers are beginning to rework curriculum continua, materials and staff development activities with the intention of increasing student learning through improved instruction, and measuring that with districtwide testing programs. In the fourth district, a fundamental school philosophy has produced a mix of whole group instruction, grade level standards, retention in grade, corporal punishment as methods for improving student skills in reading, writing and mathematics with measurement assessed by districtwide tests.

A contributing factor in the relative clarity of district goals is that the central office administrators in each of the four districts are a relatively stable group. In spite of changes at the Superintendent level, the individuals responsible for curriculum, instruction, and supervision of elementary and secondary levels have worked together over a long period of time. In all four districts these individuals have evolved methods of communicating and resolving difficulties. This stability, rather than leading to stagnation, seems to have contributed in three of the four districts a sense of direction, more coherent than

one would have thought possible. Partly this is due to the previously mentioned commitment of the staffs to instructional improvement; partly, it may be due to a general high degree of competence at the district level. In these three districts respondents, in commenting on the work of their colleagues, frequently alluded to the high quality of their technical and interpersonal skills.

ABOUT DISTRICTWIDE TESTING

Types of tests administered. All four districts adminstered the following tests: district-developed criterion-referenced tests; district-developed proficiency examinations; the norm-referenced CTBS test either for Title I students only or for all students at specific grade levels; State Assessment Program tests.

Ways in which tests are linked to instruction. The four districts try to link tests with instruction in a number of ways. In one district, district-developed criterion-referenced tests are the hub around which curriculum, classroom management, materials, staff development are organized. In this district, the other three types of tests are only to satisfy federal or state reporting requirements. Although student scores on these other tests are reviewed within the district, the degree to which this analysis modifies instruction is minimal.

In the second district, the courts have mandated that CTBS scores be used as a measure of student achievement and, indirectly, of the district's committment to providing equal educational opportunity. Therefore, the district is attempting to find ways of raising student scores; this may mean abandoning their existing curriculum-relevant CRTs in order to modify the curriculum in the direction of CTBS content.

In the third district, the Board has also focused on CTBS scores as the particular measure of district success in increasing student learning. The Board is concerned that children in this predominantly blue-collar district do as well as students in other districts. In spite of explanations by the test director as to why CTBS tests may not be an adequate reflection of student learning in this district, the Board persists. Concurrently, this district is trying to develop a CRT system for internal management. They hope that the fundamental philosophy, combined with district CRT-based accountability procedures, will lead to an increase in CTBS scores.

In the fourth district, there is as yet no single test to which the district has anchored its instructional improvement effort. Scores from the State Assessment Program tests are being used to identify areas where curriculum might be augmented or teaching strategies fine-tuned. The district proficiency examinations --not yet integrated with a nascent CRT system-- are stimulating the development of remedial programs, new staff development activities, and a review of course offerings in the secondary schools.

It is interesting to note that districts' domestication characteristics are in varying states of transition. While the future funding and supply of students in all the districts are not tied to the test results, some districts are finding that poor test results may precipitate externally mandated changes such as forced busing or internally directed (school board) changes in personnel. It may be that as boundary permeability increases, goal diffusion decreases. This might happen, however, only when the pressures on a district are in a consistent direction.

Characteristics of tests that motivate linkage. In each of our four districts efforts were being made to link one or another type of test with instruction. The district which seemed to have made the most progress in this direction was the one which selected a single type of test --their own CRT system-- as the anchor for such a linkage. The three other districts were attempting to use more than one type of test to analyze their curriculum or identify areas of student deficiency. In these districts, district administrators were realizing that the different assumptions underlying the construction of the tests implied different courses of action.

The impetus for linking norm-referenced test scores with instruction came from the court in one case, and the School Board in another. Comparability of students' performance with those in other schools or districts, was the primary characteristic of these tests which attracted the court and the Board. The two districts under pressure to use CTBS scores as a measure of student achievement must do so to prove that their students are becoming the equal of students in other places. However, the most salient characteristics of test for district administrators who think about linking testing and evaluation with instruction, was the test's compatibility with existing curriculum and instruction. This moved the districts towards criterion-referenced tests.

ABOUT DISTRICTWIDE EVALUATION

Types of evaluations administered. Districts differed from one another in the extent to which they utilized evaluation information. Often, respondents spoke about evaluation as if it were synonymous with testing. The distinctly evaluative activities linked with instruction were of two kinds --the first associated with the Program Quality Review teams, and the second associated with Consolidated Application Forms.

Ways in which evaluations are linked to instruction. In each of the four districts some evaluations, defined as distinct from testing, were occurring. In all districts, Title I requirements were being met --usually through the provision of technical assistance by a central office testing person-- by school administrators. In the large decentralized district, completing the Consolidated Application Form was elevated into an important school-site planning activity. There was much central office effort directed to assisting schools with needs assessment, planning activities and monitoring activities.

Each of the districts was visited by Program Quality Review teams from the state, who were evaluating school programs funded by the School Improvement Program. Two districts adopted this form of evaluation internally and conducted mock site visits in advance of the state team.

In only the small centralized district did central office staff annually survey parents and students about their attitudes towards work and specific subject areas.

The small elementary district takes very seriously the visits of the Program Quality Review team. It is our impression that principals' and teachers' activities are shaped by the anticipation of those visits. Since the PQR teams assess implementation of school-site plans, and since this district is committed to an instructional approach very similar to that endorsed by the School Improvement Program, it does not seem unusual for teachers to take this evaluation seriously. They appear to regard the visits of the PQR teams as a form of self-evaluation.

The staff of the Evaluation Office in the large urban district, as noted earlier, provides technical assistance to schools in preparing

Consolidated Application Forms. Information in these forms is intended to be the core of the school's annual improvement plan. Regular visits of district evaluators to report back to the school on student progress as measured by tests, are supposed to modify classroom activities.

In the third district, evaluations conducted in connection with federal and state funds were used only to provide data for the following year's applications. Although district officials acknowledged that information from these evaluations might prove helpful to the program, they had little inclination to make this activity a top priority.

ABOUT T/E/I LINKING SUBSYTEMS

From our interview data, we have described a construct called the T/E/I linking subsystem. This subsystem does not exist on an organizational chart in any of our districts. Only in the large urban district is there an evaluation unit. In each of the other districts there are one or two individuals charged with responsibility for supervising testing and evaluation. However, in each of the districts some or all of the elements constituting a subsystem are present.

The districts differ from one another in the extent to which they conceptualize what they are doing as a subsystem. In the small elementary district there is a very clear and comprehensive view of operations and coordinating mechanisms supporting the teaching/testing/teaching cycle.

In the other three districts, respondents discuss what they are doing in terms of managing the linkage between testing and evaluation and instruction in an incomplete or narrow manner.

In the largest district, each individual sees the T/E/I susbystem primarily from his or her own job responsibilities. Although the district has a number of coordinating mechanisms, the size of the district precludes overlapping of responsibilities and informal conversations characteristic of the smaller district. The two medium-size districts differ from one another on the amount of agreement underlying their approach. In one district, staff development is the agreed-upon focus. Here there is some discussion about how to connect testing and evaluation data with the commitment to teacher decision-making. In the other district, there is ferment surrounding the introduction of fundamental schools. Disagreements on this issue spill over into disagreements on how to develop and manage testing and evaluation systems.

Thus, the districts' levels of looseness of coordination or tightness of coupling can be thought of as being arrayed along a continuum --with one tightly-coupled district at one end and two that have some coupling mechanisms in place and one which has made little progress, to date, in coordinating their activities.

Each of the districts is at a different phase in the development of a T/E/I linking subsystem. The large urban district has had for some time a complex arrangement for coordinating their decentralized system. This arrangement has been smothly operating for many years but is only now being made to carry the load of linking CTBS tests with instruction. The small elementary district has been moving toward a highly centralized coordinated system for eight years. They now regard the T/E/I linking subsystem as being in place. Strains are being put on it due to expansion of the district to include secondary schools. One of the remaining districts has been working at staff development for five years

and is only now beginning to think of testing outcomes relevant to the context of staff development. The last district, impelled by their transition to fundamental schools, is in only in the first stage of developing a linking subsystem.

The districts are therefore in different phases in relation to one another. Furthermore, the origins of their subsystems are quite different --one is clearly coming from an emphasis on school-site autonomy, one from an emphasis on teacher capacity-building, one from an emphasis on individualized learning, and one from an emphasis on fundamental education.

What the districts do share in common is the length of time it appears to take to evolve such a system. For none of the districts did the evolution of a T/E/I linking subsystem appear to be planned programmatic effort. Rather, it emerged as a series of responses to particular problems.

This patterns seems consistant with what we know about goal diffuseness and boundary penetration as characteristics of school districts as organizations. The extent to which those responses moved in the direction of linking tests with evaluation or instruction depended, in large measure, on the vision and professional preferences of the people in leadership positions. The stability of the staff made it possible for them to act periodically, even if not consistently, on their general orientation towards instructional improvement.

One of the issues raised by the reactive, episodic, and person-specific developmental history of each of the districts' subsystems are their durability. Such systems seem to take great effort to build, but can be undone in a very short period of time. Changes in personnel or other happenings over which the district has little control are potent threats.

In interviewing Central Office administrators, Board member, parents, principals and teachers in each district, we found that four types of external environments were frequently mentioned as influential. We have termed these societal, community, professional, and intradistrict. Although respondents in all four districts commented on these four types of environments the factors within each type --for example, people, events, trends, etc.-- differed from district to district, and the relative importance of each type also differed from district to district.

In the small elementary district, the four types of environment seemed to point in the same direction: towards a well managed individualized learning approach. The community wanted this, there were no overwhelming external problems distracting district administrators' energies, the professional community could provide assistance and direction, the individuals within the districts and the schools knew one another and worked well together. Even so, with all factors favorable the T/E/I linking subsystem evolved slowly and with setbacks due to teacher resistance and to an inadequate understanding of both the technology of testing and of teaching.

In the large urban district, the forces external to the district are pressuring district administrators to move in directions they would not otherwise choose. Court pressures, combined with increasing heterogeneity among students, are moving the district towards greater centralization, towards a narrowing of the curriculum, towards less teacher autonomy --in the interests of raising student scores on a norm-referenced test.

Similar forces --although mediated by the School Board, rather than the courts-- are moving one of the medium-sized districts towards a fundamental philosophy for all schools.

The other medium sized district is modifying its current practices primarily because of the state required student proficiency requirements. Anticipated community protests against the denial of diplomas to failing students is propelling the district in a direction it might not have chosen to go.

IMPACT ON CLASSROOM ACTIVITIES

In the small elementary district, teachers' behaviors seemed influenced by district policies. Teachers asserted that their teaching strategies, their use of materials, their management of the classroom had been influenced by district-level activities. In this district, the teaching/testing/teaching system made teachers feel professional, feel proud of student learning; they worked hard but felt supported by the district.

In the three other districts, there were widely different levels of awareness among teachers as to what district intentions were in relation to testing, evaluation and instruction. Teachers who had participated in the district's staff development activities acknowledged that they had been influenced in a general way by these activities and that their teaching was more targeted. The optimistic view in the district is that these teachers are becoming more able to use the diagnosis provided by student test scores because they have a number of prescriptions in their repertoire. Clinical supervision by principals to support this hope is being instituted this year.

Teachers in the fundamental district expressed a wait-and-see attitude tinged by some anxiety that tests would become a tool of teacher assessment. In this district, informal and interpersonal sanctions may be used to overcome traditionally autonomous teacher functioning.

In the large urban district, the decentralized approach to teaching with its acknowledgement of teacher autonomy may be giving way to more centralized directions. The district is preparing for this by organizing committees and working groups to encourage teacher participation and commitment.

PART III.

REFERENCES

REFERENCES

Abramowitz, S., and Tenenbaum, E., with T.E. Deal and E.A. Stackhouse. High School '77: a survey of public secondary school principals. Washington, DC: National Institute of Education, 1978.

Alkin, M.C., Daillak, R., White, P. Using evaluations: Does evaluation make a difference? Beverly Hills, Calif.: Sage Publication, 1979.

Alkin, M.C., Kosecoff, J., Fitz-Gibbon, C., and Seligman, R. Evaluation and Decision Making: The Title VII Experience. CSE Monograph Series in Evaluation (No. 4) Los Angeles: Center for the Study of Evaluation, University of California, Los Angeles, 1974.

Allison, Graham. Essence of Decision. Boston: Little, Brown, 1971.

Anderson, J.G. Bureaucracy in education. Baltimore: Johns Hopkins Press, 1968.

Bank, A. & Williams, R.C. Evaluation design project: School district organization study. Progress Report, February 1980 (Revised 3/12/80).

Barnard, C. The Functions of the executive. Cambridge, Mass: Harvard University Press, 1938.

Bentzen, M.M., et al. Changing schools: the magic feather principle. New York: McGraw-Hill, 1974.

Berman, P. Toward an implementation paradigm of educational change. Paper prepared for Program on Research and Educational Practice, National Institute of Education, 1980.

Berman, P., and McLaughlin, M.W. Federal programs supporting educational change, Vol. I: A model of educational change. Santa Monica, CA: Rand Corp., 1974. R-1589/1-HEW.

Berman, P., and McLaughlin, M.W. Federal programs supporting educational change, Vol. VIII: Implementing and sustaining innovations. R-1589/8-HEW. Santa Monica, CA: Rand, 1978.

Berman, P. and McLaughlin, M.W. An exploratory study of school district adaption. R-2010 NIE. Santa Monica, CA: Rand, 1979.

Bidwell, C.E. The school as a formal organization. Chapter 23, in Handbook of organizations. Chicago: Rand McNally & Co., 1965.

Blau, P. & Scott, W.R. Formal organisations. San Francisco: Chandler Publishing, 1965.

Boruch, Robert F. and Cordray, David S., Report on use of evaluations at federal, state and local levels, Sept. 1980, (unpublished).

Boyd, T., McKenna, B.H., Stoke, R.E. & Yadunsky, J. A study of testing priorities in the Royal Oaks (Michigan) public schools. Royal Oaks, Michigan: Royal Oaks City School District, 1975. (ERIC Document Reproduction Service No. ED 117 161)

Boyd, W.L. The changing politics of curriculum policy-making for American schools. Review of Educational Research, 48 (4): 577-628, 1978.

Boyd, W.L. The politics of curriculum change and stability. Educational Researcher, 8 (2): 12-18, 1979 (a).

Boyd, W.L. Retrenchment in American education: the politics of efficiency. Paper read at American Educational Research Association meetings, 1979 (b).

Burns, T., and Stalker, G.M. The management of innovation. London: Tavistock Publications, 1961.

Carlson, R.O. Executive successors and organizational change. Chicago: University of Chicago, Midwest Administration Center, 1962.

Carlson, R.O. Environmental constraints and organizational consequences: the public school and its clients. In D.E. Griffiths (Ed.) Behavioral science and educational administration. 63rd Yearbook, National Society for the Study of Education. Chicago: University of Chicago Press, 1964. Pp. 262-76. (Also reprinted in Baldridge, Deal & Ancell, 1975.)

Clark, D.L., (ed) "In consideration of Goal-Free Planning: The Failure of Traditional Planning Systems in Education", New Perspectives on Planning in Educational Operations, July 1980 (unpublished).

Cohen, M., March, J., & Olsen, J. A garbage can model of organizational choice. Administrative Science Quarterly, 1972, 17, 1-25.

Corwin, R.G. Professional persons in public organizations. Educational Administration Quarterly, 1:1-22, 1965.

Corwin, R.G. The school as a formal organization. In L.C. Deighton (Ed.) Encyclopedia of Education. New York: Macmillan and Free Press, 1971. Vol. 8.

Corwin, R.G. Models of educational organization. Chapter 7 in Review of research in education. New York: American Educational Research Association, 1974 (pp. 247-287).

Cromer, S., and Spakter, M. A scale to assess attitudes towards aptitude testing. Measurement and evaluation in guidance. 1968, I, (2)

CSE Long Range Plan, 1978-1982, in response to official solicitation of long range plans of labs and centers by NIE, March 25, 1977, submitted September 30, 1977.

CSE Three Year Plan proposal to the National Institute of Education under a long term special institutional relationship. December 1979- November 1982. The Evaluation Design Project, revised 8/30/79.

Daft, R.L. and Becker, S.W. The innovative organization: innovation adoption in school organizations. New York: Elsevier, 1978.

Daillack, R.H. Evaluators at Work, Center for the Study of Evaluation, UCLA, September 1980.

Dalton, M. Men who manage. New York: Wiley & Sons, 1959

David, J. Local uses of Title I evaluation. Palo Alto, California: Stanford Educational Policy Research Center, 1978.

Deal, T.E. Linkage and information use in educational organizations. Paper read to American Educational Research Association meetings, 1979.

Downs, G.W., Jr., and Mohr, L.B. Conceptual issues in the study of innovation. Administrative Science Quaterly, 21: 700-714, 1976.

Elboim-Dror, R. Some characteristics of the education policy formation system. Policy Sciences, 1:231-253, 1970.

Elboim-Dror, R. Organizational characteristics of the educational system. Journal of Educational Administration, 9, May, 1973.

Etzioni, A. A comparative analysis of complex organizations. New York: Free Press, 1961.

Farrar, E., DeSanctis, J. and Cohen, D.K. Views from below: implementation research in education, Cambridge, MA: The Huron Institute, 1979.

Firestone, W.A. Images of schools and strategies for change. Philadelphia: Research for Better Schools, Inc., 1978.

Firestone, W.A. and Herriott, W.A. Images of the school: an exploration of the social organization of elementary, junior high and high schools. Philadelphia, PA: Research for Better Schools, 1980.

Gallaher, A. Directed change in formal organizations: the school system. In R.O. Carlson (Ed.) Change processes in the public schools. Eugene, Oregon: CASEA, University of Oregon, 1965.

Goodlad, J. The Dynamics of Educational Change: Towards Responsive Schools. New York: McGraw-Hill, 1975.

Goodlad, J. What Schools are For, Bloomington Indiana: Phi Delta Kappa Educational Foundation, 1979

Goodlad, J. and Klein, M.F. Behind the classroom door. Belmont, CA: Wadsworth Publishing Co., 1970.

Gorth, W.P., and Perkins, M.R. A Study of minimum competency testing program final summary and analysis report, December 1979.

Goslin, D.A. The use of standardized tests in elementary schools. New York: Russell Sage Foundation, 1965.

Gouldner, A. Patterns of industrial bureaucracy. New York: The Free Press, 1954.

Guba, E.A. Toward a Methodology of Naturalistic Inquiry in Educational Evaluation. CSE Monograph No. 8, Center for the Study of Evaluation, Los Angeles, 1978.

Halperin, S. The Educational Arena, Educational Evaluation and Policy Analysis. 2(1):27-36, 1980.

Hanson, E.M. Educational administration and organizational behavior. Boston: Allyn & Bacon, 1979.

Herriott, R.E. and Hodgkins, B.J. The environment of schooling: Formal education as an open social system. Englewood Cliffs, New Jersey: Prentice-Hall, 1973.

House, E. Three perspectives on innovation: The technological, the political, and the cultural. Paper prepared for Program on Research and Educational Practice, National Institute of Education, 1980.

Hoy, W. and Miskel, C. Educational administration: Theory, research and practice. New York: Random House, 1978.

Katz, D. and Kahn, R. The social psychology of organizations. New York: Wiley, 1966.

Katz, D. and Kahn, R. The social psychology of organizations (2nd ed.) New York: Wiley & Sons, 1978.

Kauffman, T.D. State assessment programs: Current studies and a look ahead. Paper presented at the Annual Meeting of the National Assessment of Educational Programs, Denver, Colorado, 1979.

Kirkland, M. The effects of tests on students and schools. Review of Educational Research, 1971, 41 (4): 303-350.

Kirst, M.W. Interview, Educational Evaluation and Policy Analysis, 2 (1): 77-83, 1980.

Lawrence, P. and Lorsch, J. Organization and environment: Management differentiation and integration. Cambridge, Mass.: Harvard University Graduate School of Business Administration, 1967.

Lieberman, A. Political and economic stress and the social reality of schools. Teachers College Record, 79 (2): 259-267, 1977.

Lindbloom, C.E. The Science of Muddling Through, Public Administration Review. 19 (Spring 1959).

Lortie, D.C. School teacher. Chicago, Ill.: University of Chicago Press, 1975.

Lortie, D.C. An exploration of urban school structure and teacher professionalism. Chicago: Center for New Schools, 1977(a).

Lyon, C.D., Doscher, L., McGranahan, P., and Williams, R. Evaluation and school districts. Los Angeles: Center for the Study of Evaluation, UCLA Graduate School of Education, 1978.

March, J.G. Footnotes to Organizational Change, Project Report No. 80-A6, Institute for Research in Educational Finance and Governance, Stanford University, 1980.

March, J.G., and Simon, H.A. Organizations. New York, Wiley, 1958.

McLaughlin, M.W. Evaluation and reform: ESEA Title I, 1965. A Rand Educational Policy Study. Cambridge, Mass.: Bollinger Publishing Co., 1975.

Meyer, J.W. Research on school and district organization. Paper presented at Annual Sociology of Education Association Conference, San Diego, 1977.

Meyer, J.W., and Rowan, B. Institutionalized organizations: Formal structure as myth and ceremony. American Journal of Sociology, September, 1977, pp. 340-363.

Meyer, J.W., and Rowan, B. The structure of educational organizations. In M.W. Meyer and Associates (Eds.). Environments and organizations. San Francisco: Jossey-Bass, 1978. Pp. 78-107.

Miles, M.B. Common Properties of Schools in Context: The Backdrop for Knowledge Utilization and School Improvement. Paper prepared for Program on Research and Educational Practice, National Institute for Education, 1980.

Moeller, G.H., and Charters, W.W. Relationship of bureaucratization to a sense of power among teachers. Administrative Science Quarterly, March, 1966, 10 (4).

Niedermeyer, F.C. A model for the implementation of outcomes-based instructional products. Paper read at American Educational Research Association meetings, 1979.

O'Shea, D. School District Evaluation Efforts: Contradictions and Implications. In Faculty Papers Dealing With School District Research, Development and Evaluation Units. Center for the Study of Evaluation, 1980.

Owens, R.G. Organizational behavior in schools. Englewood Cliffs, New Jersey: Prentice-Hall, 1970.

Packard, J.S. Schools as work organizations. Paper read to Symposium on Organizational Characteristics of Schools, University of California at Santa Barbara, February 10, 1977. Eugene, OR: Center for Educational Policy and Management, University of Oregon, 1977.

Patton, M.Q. Utilization-focused evaluation. Beverly Hills, CA.: Sage Publications, Inc., 1978.

Perrow, C. Complex organizations: A critical essay. Glenview, Illinois: Scott Foresman, 1979.

Pincus, J. Incentives for innovation in the public schools. Review of Educational Research, 44:113-144, 1974.

Pincus, J., and Williams, R.C. Planned change in urban school districts. Phi Delta Kappa, June 1979, pp. 729-733.

Rogers, E.M., and Rogers-Agarwala, R.R. Communications in organizations. New York: The Free Press, 1976.

Sarason, S.B. The culture of the school and the problem of change. Boston: Allyn & Bacon, 1971.

Sieber, S.D. Incentives and disincentives for knowledge utilization in public education. Paper prepared for Program on Research and Educational Practice, National Institute of Education, 1980.

Spence, L.D., Takei, Y., and Sim, F.M. Conceptualizing loose coupling: Believing is seeing or the garbage can as myth and ceremony. Paper read to American Sociological Association meeting, 1978.

Testing, Assessment and Evaluation Unit. Research area plan in program development and evaluation. Draft, February 1980.

Thompson, J.A. Organizations in action. New York: McGraw-Hill Book Co., 1967.

Tomblin, E., and Harding, E. Small School District Utilization of Program Evaluation: A Survey of Board Members and Superintendents. Paper presented at the American Educational Research Association annual meeting, Boston, Massachusetts, April 1980.

Tyler, R.W., and White, S.H. (Chairman). Testing, teaching and learning: Report of a conference on research on testing. Washington: National Institute of Education, 1979.

Weber, M. The theory of social and economic organization. A. M. Henderson and T. Parson (Trans. & Ed.). Oxford University Press, 1974.

Weick, K.E. Educational organizations as loosely coupled systems. Administrative Science Quarterly. 1976, 21, 1-19.

Weinheimer, A. Critical needs for research and development in evaluation: Observations from a federal agency's perspective. Prepared for the American Educational Research Association session in 21.01, April 10,1980.

Weisbord, M. Input- versus output-focused organizations: notes on a contingency theory of practice. In W.W. Burke (Ed.) The cutting edge: current theory and practice in organization development. La Jolla, CA: University Associates, 1978. Pp. 13-26.

Weiss, C.H. (Ed.) Evaluating action programs. Boston: Allyn and Bacon, 1972.

Williams, R. (Ed) Faculty papers dealing with school district research, development and evaluation trends. Center for the Study of Evaluation, UCLA Graduate School of Education, 1979.

Wirt, Fred M. and Michael W. Kirst. The Political Web of American Schools. Boston: Little, Brown, 1972.

Wise, A.E. The hyper-rationalization of American education. New York University Education Quarterly, 8(4), 2-6,1977.

Yeh, J.P. Test use in schools, Los Angeles, California: Center for the Study of Evaluation, 1978.

Zucker, L.G. Institutional structure and organizational processes: the role of evaluation units in schools. In Faculty Papers Dealing with School District Research, Development and Evaluation Units, Center for the Study of Evaluation, 1980.

APPENDICES

QUESTIONNAIRE MODULES

INTRODUCTION

Let me start by telling you about the purpose of this interview. I'm _____ from UCLA. I work at a research institution in the Graduate School of Education, called the Center for the Study of Evaluation. The Center is one of 17 labs and centers funded by the National Institute of Education to do research and development in education. CSE has been doing this since 1965.

Currently, the Center is interested in two broad research areas--testing and evaluation. We've just begun a project to find out how school districts conduct testing or evaluation programs. That's why we're here in the _____ district.

I want to assure you that our talk is confidential; the answers we get from you and other people in your school district will be combined into an overview. Where examples and quotations are used as evidence, people and places will not be identified.

So, the purpose of our interview is to help us understand the evaluation and testing program in your district. Any questions before we begin?

I'd like to get your permission to tape-record our talk, so I don't miss anything in my notes or change your words or ideas. If you want the recorder off at any time, just press this button. (Demonstrate)

MODULE # _____

COMMENTS

DISTRICT _____

NAME _____

TITLE _____

PHONE _____

INTERVIEWER _____

DATE _____

TIME BEGUN _____ TIME ENDED _____

VISIT #_____

MODULE 1

BACKGROUND OF RESPONDENT

It will be helpful to know something about your job here and about your experience.

1.1 YOUR OFFICIAL TITLE IS . . . ?

1.2 AND HOW LONG HAVE YOU BEEN IN YOUR PRESENT POSITION?

1.3 WHICH AMONG YOUR PREVIOUS JOBS, EDUCATIONAL AND TRAINING EXPERIENCES DO YOU FIND MOST USEFUL IN DOING THE WORK YOU DO NOW, IN YOUR PRESENT POSITION?

(Probe for how, why.)

(Probe once or twice: any other background you find that you draw on in your present work?)

1.4 IF YOU WERE TO WRITE A BRIEF PARAGRAPH OR TWO AS A JOB DESCRIPTION FOR YOUR CURRENT POSITION, WHAT WOULD IT SAY?

(Don't probe: Encourage response in format called for.)

1.5 I'D LIKE AN OVERVIEW OF WHAT YOU ACTUALLY DO IN YOUR JOB OVER THE COURSE OF A YEAR. HOW ABOUT IF YOU PICK A POINT IN THE YEAR THAT MAKES SENSE TO BEGIN WITH, THEN PROCEED THROUGH A YEAR FROM THERE, TELLING ME WHAT I'D SEE YOU DOING IF I WERE HERE.

(Suggest September if respondent has trouble choosing. Walk respondent through a full year. Probe: Anything else you do at that time? Can you be more specific about that? Tell me about that in a bit more detail, please.)

1.6 LET ME JUST CLARIFY WHO YOU ARE TO REPORT TO --THAT IS, IN THE OFFICIAL LINE OF AUTHORITY.

1.7 HOW WOULD YOU CHARACTERIZE THE RELATIONSHIP BETWEEN HIS/HER POSITION AND YOURS? DOES THAT PERSON DIRECT YOUR WORK AND SUPERVISE YOU CLOSELY?

1.8 DO YOU DELIVER FORMAL REPORTS TO HIM/HER? DO YOU WORK COLLEAGUELY AND SHARE RESPONSIBILITY?

1.9 THEN, WHO ELSE IN THE DISTRICT DO YOU FIND YOURSELF TALKING TO, OR WORKING WITH, MOST OFTEN?

(Probe for purposes of talk, work done jointly, etc.)

MODULE 2

DISTRICT TESTING ACTIVITIES

Next, I'd like for us to turn to the activities here in _____ _____ directly relating to testing.

2.1 TO GET AN OVERVIEW OF TESTING HERE, I'D LIKE TO HAVE YOU HELP ME FILL IN AS MUCH OF THIS FORM AS YOU CAN (show form). WE'RE INTERESTED IN SOME VERY BASIC INFORMATION ABOUT THE TESTS ADMINISTERED, HOW OFTEN THEY'RE GIVEN, AND SO ON.

WHY DON'T WE JUST GO OVER THIS TOGETHER.

COMPLETE CHART 2.1 - TESTING

CHART 2.1 TESTING

Name of Test	How long has district used test?	Time of Administration/ Time of Results	To Whom Administered	General Purpose	Mandated, Selected or Developed	Source of Funding

2.2 THERE'S TALK TODAY ABOUT HOW CERTAIN TESTS MAY NOT BE APPROPRIATE FOR MINORITY STUDENTS --FOR BLACKS, FOR THOSE WHO ARE NOT NATIVE SPEAKERS OF ENGLISH, AND SO ON. DID THIS COME UP WHEN YOU WERE CHOOSING/DEVELOPING ANY OF THE TESTS YOU MENTIONED? (Probe for details, criticisms, reasoning.)

2.3 AT THIS POINT I'D LIKE YOU TO HELP ME UNDERSTAND MORE SPECIFICALLY HOW THE RESULTS OF THE TESTS THAT YOU ADMINISTER ARE HANDLED AND HOW THEY'RE USED. I'M INTERESTED IN HOW THEY'RE SCORED AND ANALYZED, WHO SEES THE RESULTS, AND WHAT HAPPENS THEN.

a) WHAT TESTS DO YOU THINK THAT YOUR DISTRICT PAYS THE MOST ATTENTION TO?

b) TELL ME WHAT HAPPENS WITH EACH TEST STEP-BY-STEP, FROM THE MOMENT THE STUDENTS COMPLETE THEM.

(Probe for answers to the following:

- ARE SCORES BROKEN OUT BY SEX OR ETHNICITY? WHAT FORM(S) ARE THE SCORES REPORTED IN?
- WHO ANALYZES THE SCORES, RESULTS?
- WHO FIRST SEES THE RESULTS? WHAT ROUTE DO THEY TRAVEL THROUGH DISTRICT?
- WHAT DO THOSE WHO RECEIVE THE RESULTS DO WITH THEM? (Get concrete description. Examples.)

2.4 IN SOME DISTRICTS, WE'VE BEEN TOLD PRINCIPALS AND TEACHERS DO NOT HAVE THE SKILLS TO USE TEST RESULTS. DO YOU FEEL THAT THIS IS A PROBLEM IN YOUR DISTRICT?

2.5 (If it is a problem), WHAT IS THE DISTRICT DOING TO SOLVE THE PROBLEM?

2.6 WHAT EVIDENCE DO YOU HAVE OF PRINCIPALS AND TEACHERS USING TEST RESULTS?

2.7 I AM INTERESTED IN KNOWING IF YOU RECEIVE ADVICE OR ASSISTANCE FROM OUTSIDERS FOR YOUR TESTING ACTIVITIES.

COMPLETE CHART 2.6 - OUTSIDE ASSISTANCE

CHART 2.7 - OUTSIDE ASSISTANCE

	Title/Person	Testing Program(s)	Why Contacted	Kind of Assistance	Impact
Federal					
State					
County					
University					
Consultant					
Other					

MODULE 3
DISTRICT EVALUATION AND RESEARCH

I am interested in what evaluation and research activities are performed in your district other than testing.

3.1 I'D LIKE YOU TO TELL ME ABOUT THE MANDATED EVALUATIONS YOU DO IN THIS DISTRICT.

SEE CHART 3.1 - MANDATED EVALUATIONS

CHART 3.1 - MANDATED EVALUATIONS

Program/Activity Evaluated	Mandated By	Frequency	Procedures	How Much Time Does It Take	Form of results

3.2 WHAT DO YOU DO WITH THE RESULTS FROM THESE MANDATED EVALUATIONS?

(Probe for: how district personnel use results, who else uses results, people involved, etc. Get examples.)

3.3 NOW, I'D LIKE TO KNOW IF THERE ARE ANY EVALUATION OR RESEARCH ACTIVITIES YOU CONDUCT WHICH ARE NOT REQUIRED BY AUTHORITIES OUTSIDE THE DISTRICT. THIS WOULD INCLUDE ANY ACTIVITIES THE DISTRICT OR SCHOOLS MAY CONDUCT TO GET INFORMATION ABOUT WHY OR HOW THINGS ARE GOING.

COMPLETE CHART 3.3 - NONMANDATED EVALUATIONS OR RESEARCH

I would like to have copies of any recent evaluation and research reports that are available.

CHART 3.3 - NON-MANDATED EVALUATIONS OR RESEARCH

Program/Activity Evaluated or Researched	Reason for Evaluation	Who Conducts Work (Title/ Position)	Frequency/ Origin	Procedures/ Time Involved	Form of results	Funding Source/ Amount

3.4 WHAT DO YOU DO WITH THE RESULTS FROM THESE NONMANDATED EVALUATIONS? (Probe for: how district personnel use results, who else uses results, people involved, etc. Get examples.)

3.5 SELECT ONE OF THE EVALUATIONS THAT WE'VE TALKED ABOUT AND TELL ME WHAT YOU (OR YOUR UNIT) DO DURING AND AFTER THE EVALUATION.

3.6 HOW DOES WHAT YOU HAVE JUST DESCRIBED COMPARE WITH THE WAY YOU HANDLE OTHER EVALUATIONS?

3.7 I AM INTERESTED IF YOU GET ANY ADVICE OR ASSISTANCE FROM OUTSIDERS ON YOUR EVALUATION OR RESEARCH ACTIVITIES.

COMPLETE CHART 3.7 - OUTSIDE ASSISTANCE

CHART 3.7 - OUTSIDE ASSISTANCE

	Person/Title	Which Evaluation or Research Activity	Why Contacted	Kind of Assistance Provided	Impact
Federal					
State					
County					
Inside Dist/Other Than Evaluators					
University					
Consultant					
Private Business					
Other					

201

202

MODULE 4

DISTRICT
INSTRUCTIONAL PROGRAM

At this point, I'd like to discuss the curriculum and instruction program here in (the district).

4.1 BRIEFLY, COULD YOU DESCRIBE THE CURRICULUM AND INSTRUCTION STAFF AND THEIR RESPONSIBILITIES?

4.2 DO YOU HAVE ANY SPECIAL SCHOOLS OR ALTERNATIVE SCHOOLS IN YOUR DISTRICT? PLEASE DESCRIBE THEM.

4.3 WHAT DO THE PRINCIPALS IN THIS DISTRICT DO IN RELATION TO CURRICULUM AND INSTRUCTION?

4.4 ARE PRINCIPALS GIVEN SPECIFIC DISTRICT-LEVEL DIRECTION IN FUNCTIONING AS INSTRUCTIONAL LEADERS?

4.5 IS INSTRUCTIONAL TRAINING OR ASSISTANCE AVAILABLE TO BUILDING PRINCIPALS?

4.6 ARE THERE OTHERS YOU PERCEIVE AS INSTRUCTIONAL LEADERS AT THE LOCAL SCHOOL LEVEL?

Next, I would like to talk about major emphases in instruction and content. Let's begin by talking about content.

4.7 WHAT DO YOU SEE AS THE MAJOR AREAS OF CONTENT EMPHASIS IN (NAME OF DISTRICT).

4.8 HOW DID THE DISTRICT COME TO CHOOSE THESE MAJOR AREAS OF EMPHASIS? (Probe for process: who involved, basis for decisions, evidence used, controversy surrounding choice.)

4.9 DO YOU USE STAFF DEVELOPMENT AS A PART OF THIS EFFORT? HOW?

4.10 ARE THERE PARTICULAR TEACHING STRATEGIES, OR TECHNIQUES, THE DISTRICT IS WORKING TO SEE IMPLEMENTED?

4.11 HOW DID THE DISTRICT COME TO CHOOSE THESE TEACHING STRATEGIES OR TECHNIQUES?

(Probe for process: who involved, basis for decisions, evidence used, controversy surrounding choice.)

4.12 Do you use staff development as part of the effort? How?

WHO IN YOUR DISTRICT PARTICIPATES IN MAKING DECISIONS ABOUT EACH OF THE FOLLOWING, AND WHAT INFORMATION DO THEY USE: TEST SELECTION, STUDENT PLACEMENT, CONTINUUM DEVELOPMENT?

4.13 TEXTBOOK SELECTION?

4.14 WHO IN YOUR DISTRICT PARTICIPATES IN MAKING DECISIONS ABOUT STUDENT PLACEMENT, AND WHAT INFORMATION DO THEY USE?

4.15 WHO IN YOUR DISTRICT PARTICIPATES IN MAKING DECISIONS ABOUT CONTINUUM DEVELOPMENT, AND WHAT INFORMATION DO THEY USE?

If tests or evaluations have not been mentioned, ask

4.16 DOES INFORMATION FROM THE TESTING PROGRAM OR FROM EVALUATION FINDINGS INFLUENCE ANY OF THESE?

MODULE 5

ATTITUDES TOWARDS TESTING AND EVALUATION

I'd like to get a sense of how people in your district view your district's testing/evaluation activities.

5.1 IN MANY PLACES, TEACHERS ARE NOT POSITIVE ABOUT DISTRICT EVALUATION/ TESTING. IN OTHERS THEY SEEM TO BE QUITE SUPPORTIVE. CAN YOU CHARACTERIZE HOW YOUR TEACHERS FEEL?

(Probe for: differences among teachers by grade level, school, by experience, by time in district.)

HAVE TEACHERS' OUTLOOKS ON TESTING CHANGED OVER TIME? WHEN? HOW? WHAT MIGHT ACCOUNT FOR CHANGES?

ARE THERE SOME TESTING PROGRAMS ABOUT WHICH TEACHERS FEEL MORE POSITIVELY, NEGATIVELY? IF SO, WHICH? WHAT SEEMS TO ACCOUNT FOR THESE DIFFERENCES?

5.2 WE'VE BEEN TALKING ABOUT TEACHERS' VIEWPOINTS TOWARDS TESTING AND EVALUATION. HOW ABOUT THOSE OF PRINCIPALS?

(Probe for: differences among principals by school, by experience, by time in district.)

HAVE PRINCIPALS' OUTLOOKS ON TESTING CHANGED OVER TIME? WHEN? HOW? WHAT MIGHT ACCOUNT FOR CHANGES?

ARE THERE SOME TESTING PROGRAMS ABOUT WHICH PRINCIPALS FEEL MORE POSITIVELY, NEGATIVELY? IF SO, WHICH? WHAT SEEMS TO ACCOUNT FOR THESE DIFFERENCES?

5.3 NOW, COULD YOU TALK FOR A FEW MOMENTS ABOUT WHETHER THE BOARD OF EDUCATION HAS PARTICULAR CONCERNS ABOUT TESTING AND EVALUATION?

(Probe for: changes in attitudes; impact of attitudes.)

5.4 ONCE AGAIN, WITH REGARD TO TESTING/EVALUATION IN GENERAL, WHAT ARE YOUR IMPRESSIONS OF YOUR COLLEAGUES' ATTITUDES IN THE DISTRICT OFFICE?

(Probe for: changes in attitudes; impact of attitudes.)

Finally, can you tell me about your own perspective on the district's testing program/evaluation program? In particular,

5.5 HOW DOES IT COMPARE WITH WHAT YOU'D LIKE TO BE DOING? (Obtain reasons for differences between actual and desired.) (Probe for specific details of desired program.)

5.6 IF THERE WERE TEST EVALUATIONS YOU COULD ELIMINATE, WHAT WOULD THEY BE AND WHY?

5.7 WHAT PROBLEMS ARE THERE IN REGARD TO TESTING AND EVALUATION?

MODULE 6
DISTRICT DYNAMICS

Let's look now at the district in general.

6.1 AS YOU SEE IT, WHAT ARE THE MAJOR PROBLEMS OR CONCERNS OF THE DISTRICT?

6.2 WHAT IS YOUR ASSESSMENT OF MORALE IN THE DISTRICT? WHY IS THIS SO?

6.3 DOES THE UNION (OR TEACHERS' ASSOCIATION) AFFECT IN A SIGNIFICANT WAY THE DISTRICT'S TESTING/EVALUATION AND INSTRUCTIONAL PROGRAMS? CAN YOU GIVE ME SOME EXAMPLES?

6.4 DOES THE ADMINISTRATORS' ORGANIZATION AFFECT IN A SIGNIFICANT WAY THE DISTRICT'S TESTING, EVALUATION AND INSTRUCTIONAL PROGRAMS? SPECIFIC EXAMPLES?

6.5 IN GENERAL, HOW ARE RELATIONS BETWEEN YOUR OFFICE AND OTHERS IN THE DISTRICT?

6.6 WHATEVER THE FORMAL ORGANIZATIONAL CHARTS LOOK LIKE, THERE ARE OFTEN A FEW PEOPLE WHO HAVE STRONG INFLUENCE ON WHAT GOES ON IN THE DISTRICT. IS THAT THE CASE HERE? IT WOULD BE HELPFUL TO KNOW WHO HAS THIS KIND OF INFLUENCE.

6.7 WHAT INFORMAL GROUPS EXIST IN THIS DISTRICT WHICH AFFECT DECISION-MAKING? (EXAMPLE - KITCHEN CABINET.)

6.8 AS YOU LOOK AHEAD, WHAT CHANGES DO YOU SEE AS LIKELY TO OCCUR IN THE SHORT-RANGE FUTURE? HOW DOES THIS AFFECT THE CENTRAL OFFICE AND THE SCHOOLS?

CLOSING: You've given me a good deal of your time, and I greatly appreciate it. Our talk has been an important contribution to the project we're working on jointly with your district. It has been very valuable. Thank you very much. I look forward to talking with you again.

MODULE 7
DISTRICT HISTORY

Today we're trying to find out about the recent history of the district as you have known it.

7.1 CAN YOU SUMMARIZE FOR ME THE HISTORY OF YOUR OWN RELATIONSHIP WITH THIS DISTRICT?

(Probe: When did you first become involved? In what capacity? Then what?)

Roles	From	To

7.2 I WOULD LIKE YOU TO MAKE SOME COMPARISONS BETWEEN NOW AND THE TIME WHEN YOU WERE FIRST ACQUAINTED WITH THE DISTRICT. WHAT WAS THE COMMUNITY LIKE THEN AS COMPARED WITH NOW?

7.3 HOW WAS THE CENTRAL OFFICE DIFFERENT THEN THAN NOW?

7.4 HOW WAS CLASSROOM INSTRUCTION DIFFERENT THEN THAN NOW?

7.5 IF YOU WERE WRITING A HISTORY OF THE DISTRICT, WHAT EVENTS OR TRENDS WOULD YOU SAY WERE MOST INFLUENTIAL IN MAKING SCHOOLS AND CLASSROOMS THE WAY THEY ARE TODAY IN THIS DISTRICT?

(Probe: External events? Internal events? In what way? Give examples.)

7.6 WHAT PEOPLE WERE MOST INFLUENTIAL?

(Probe: Outside the district? Inside the district? In what way? Give examples.)

7.7 CAN YOU SUMMARIZE THE HISTORY OF THE EVALUATION AND TESTING POSITIONS IN YOUR DISTRICT?

NOW, LET's TURN SPECIFICALLY TO HOW CATEGORICAL PROGRAMS, TESTING AND EVALUATION HAVE INFLUENCED CLASSROOM PRACTICE. I WOULD LIKE YOU TO TELL ME YOUR OPINION AS TO WHETHER EACH OF THE FOLLOWING HAD HIGH, MEDIUM OR LOW INFLUENCE ON THE WAY TEACHERS TEACH.

		HIGH	MED.	LOW	N/A	COMMENTS
7.8	TITLE I: PROGRAMS					
7.9	TITLE I: EVALUATIONS					
7.10	CTBS TESTING					
7.11	EARLY CHILDHOOD EDUCATION: PROGRAM					
7.12	EARLY CHILDHOOD EDUCATION: EVALUATIONS					
7.13	CALIFORNIA ASSESSMENT PROGRAM (CAP)					
7.14	TITLE IV C: PROGRAMS					
7.15	TITLE IV C: EVALUATIONS					
7.16	BILINGUAL: PROGRAMS					
7.17	BILINGUAL: EVALUATIONS					
7.18	ESAA: PROGRAMS					
7.19	ESAA: EVALUATIONS					
7.20	HART BILL PROFICIENCY TESTING					

		HIGH	MED.	LOW	N/A	COMMENTS
7.21	DIAGNOSTIC AND PRE-SCRIPTIVE TEACHING					
7.22	MASTERY LEARNING					
7.23	BACK-TO-BASICS					
7.24	TEACHER UNIONS					
7.25	COMMUNITY ATTITUDES					

7.26 FROM YOUR PERSPECTIVE, WHAT HAS THIS DISTRICT BEEN DOING TO USE EVALUATION AND TESTING TO IMPROVE INSTRUCTION?

(Probe: When did it start? Give specific examples.)

7.27 WHAT EVIDENCE DO YOU HAVE ABOUT WHETHER IT IS WORKING OR NOT?

7.28 WHAT PROBLEMS OR RESISTANCE HAS THERE BEEN?

7.29 WHAT DID THE DISTRICT DO ABOUT THESE PROBLEMS OR RESISTANCE?

7.30 FROM YOUR POINT OF VIEW, WHAT ARE THE POSITIVE EFFECTS OF EVALUATION AND TESTING REQUIREMENTS?

7.31 THE NEGATIVE EFFECTS?

7.32 ANYTHING ELSE YOU WANT TO TELL US?

THANK YOU VERY MUCH FOR YOUR TIME.

MODULE 8A

CLASSROOM IMPACT - TEACHERS

Before we begin, it would be helpful to know something of your teaching background.

8.1 HOW MANY YEARS HAVE YOU BEEN TEACHING? _____

8.2 IN THIS DISTRICT? _____

8.3 IN THIS SCHOOL? _____

8.4 WHAT GRADE LEVELS/SUBJECTS HAVE YOU TAUGHT? _____

8.5 AND YOU ARE NOW TEACHING _____ (GRADE/SUBJECTS)?

8.6 WHERE DID YOU DO YOUR TEACHER PREPARATION WORK?

8.7 WHEN? _____

8.8 AND ONE MORE THING. <u>WITHIN THE PAST TWO YEARS</u>, HAVE YOU PARTICIPATED IN ANY STAFF DEVELOPMENT OR PROFESSIONAL TRAINING EITHER INSIDE OR OUTSIDE THE DISTRICT? PLEASE DESCRIBE WHAT IT WAS AND YOUR REACTION TO IT.

<u>KIND</u>	<u>DURATION</u>	<u>REACTION</u>

8.9 COULD YOU VERY BRIEFLY DESCRIBE FOR ME HOW YOU TEACH EACH SUBJECT? (Probe for mode of instruction --total class/individuals; content; instructional materials used, etc.

Please compare what you have just told me with the way you have done things in the recent past. For example:

8.10 IS THE CONTENT DIFFERENT? IN WHAT WAY? (Get examples.) WHY DID YOU CHANGE? (Probe for specifics.)

8.11 ARE YOUR TEACHING METHODS AND MATERIALS DIFFERENT? IN WHAT WAY? (Get examples.) (Probe for specifics.)

8.12 IS THE WAY YOU THINK ABOUT CHILDREN AND HOW THEY LEARN DIFFERENT NOW THAN IT HAS BEEN? IN WHAT WAY? (Get examples.) WHY DID YOU CHANGE? (Probe for specifics.)

I'd like to quickly go through the following list. Tell me whether each topic has had a lot, some, a little, or no influence on how you have taught during this school year. Explain what you mean. (Be as detailed as time permits.)

a- a lot
b- some
c- a little
d- none

		COMMENTS
8.13	DISTRICT CONTINUUM	
8.14	DISTRICT CRT's	
8.15	TITLE I PROGRAMS	
8.16	CTBS TESTS	
8.17	CAP TESTS	
8.18	ECE/SIP PROGRAMS	
8.19	MAR (PQR) (or mock MAR) VISIT TO YOUR SCHOOL	
8.20	PRINCIPAL VISITS TO YOUR CLASSROOM	
8.21	DISTRICT STAFF DEVELOPMENT PROGRAM	
8.22	RESOURCE PEOPLE: AIDES/SPECIALISTS, ETC. (SPECIFY)	
8.23	ANY OTHER MAJOR INFLUENCE (SPECIFY)	

8.24 DO YOU THINK YOUR RESPONSES ON THE PRECEDING LIST OF TOPICS ARE TYPICAL OF OTHER TEACHERS IN THIS SCHOOL? (Probe: Differences by grade level, philosophy, question. Be as detailed as time permits.)

8.25 WHAT IS THE MOST IMPORTANT THING YOU THINK I SHOULD KNOW ABOUT THIS DISTRICT?

Thank you very much for taking the time to talk to me. I appreciate it.

MODULE **8B**

CLASSROOM IMPACT - PRINCIPALS AND OTHERS

8.1 YOUR PRESENT JOB TITLE IS? _____

8.2 HOW MANY YEARS HAVE YOU BEEN DOING THIS? _____

8.3 WHAT WERE YOU DOING BEFORE?

ROLE	FROM	TO	WHERE

8.4 WITHIN THE PAST TWO YEARS, HAVE YOU PARTICIPATED IN ANY STAFF DEVELOPMENT OR PROFESSIONAL TRAINING EITHER INSIDE OR OUTSIDE THE DISTRICT? PLEASE DESCRIBE WHAT IT WAS AND YOUR REACTION TO IT.

KIND	DURATION	REACTION

8.5 COULD YOU LIST FOR ME THE FOUR OR FIVE THINGS YOU SPEND MOST OF YOUR TIME ON DURING THE WEEK?

8.6 PLEASE DESCRIBE FOR ME WHAT YOU DO IN RELATION TO CLASSROOM INSTRUCTION.

8.7 HAVE YOUR OWN JOB RESPONSIBILITIES CHANGED RECENTLY? EXPLAIN.

8.8 HAS THE CONTENT TEACHERS COVER IN THEIR CLASSROOM CHANGED RECENTLY? EXPLAIN.

8.9 HAVE THEIR INSTRUCTIONAL METHODS CHANGED? EXPLAIN.

8.10 HAVE THEIR INSTRUCTIONAL MATERIALS CHANGED? EXPLAIN.

I'd like to quickly go through the following list. Tell me whether each topic has had a lot, some, a little, or no influence on how classroom instruction has been handled in the last year. Explain what you mean. (Be as detailed as time permits.)

a- a lot
b- some
c- a little
d- none

		COMMENTS
8.11	DISTRICT CONTINUUM	
8.12	DISTRICT CRT's	
8.13	TITLE I PROGRAMS	
8.14	CTBS TESTS	
8.15	CAP TESTS	
8.16	ECE/SIP PROGRAMS	
8.17	MAR (PQR) (or mock MAR) VISIT TO YOUR SCHOOL	
8.18	YOUR WORK IN VISITS TO CLASSROOMS	
8.19	DISTRICT STAFF DEVELOPMENT PROGRAM	
8.20	RESOURCE PEOPLE: AIDES/SPECIALISTS, ETC. (SPECIFY)	
8.21	ANY OTHER MAJOR INFLUENCE (SPECIFY)	

8.22 WHAT IS THE MOST IMPORTANT THING YOU THINK I SHOULD KNOW ABOUT THIS DISTRICT?

Thank you very much for taking the time to talk to me. I appreciate it.

PROJECT DOCUMENTS

LETTER OF SOLICITATION

February 21, 1980

Dear_____,

We are writing to ask you formally to participate with us in a project that we think will be of great interest to your district and to the Center for the Study of Evaluation. We are beginning an NIE funded study likely to benefit school districts wishing to improve their instructional programs through the use of testing and evaluation information.

As you may know, the Center for the Study of Evaluation at UCLA has been interested, over the past three years, in studying evaluation and testing activities as they are carried out in school districts. We have already completed a national survey describing the organization of evaluations. In this phase we will be collaborating with a small number of California districts to explore how evalustion and testing activities can be linked with instructional decision-making to improve education for children.

Our conversation that had with on indicated that your district shares such an interest. It appears that you have given thought for some time to how a district can manage data-based instructional improvement.

The collaborative relationship we have in mind would involve several of our CSE staff members visiting your district on several occasions between now and June. In addition, we hope that a knowledgeable person in evaluation could become a member of our Working Group to discuss the issues involved in linking testing and evaluation as well as review our findings. Since there will be representatives of five California districts on this Working Group, we anticipate the sharing of experience and information to be very stimulating. We hope to sustain the collegial Working Group relationship with your district over the next several years, although information data collected in hour district may occur only during the first year. During the second and third years, we will be interviewing in an additional eight districts.

By the end of the three-year period, we expect to have a greater understanding of how testing and evaluation data can facilitate instructional renewal. We will have identified problems common to many districts as well as the strategies that some districts have found to overcome them. We trust that our work will be beneficial to other school districts wanting to improve instruction by linkages with testing and evaluation. We may also have something of value to say to federal and state legislators and administrators who devise testing and evaluation requirements.

We expect that we need to set up interviews, primarily with central office personnel, but also with principals and teachers at both the elementary and high school level. We want to get their assistance in informing ourselves about the following resarch questions:

1. Why do school districts initiate and maintain system-supported linkages among testing/evaluation and instructional activities?

2. How have districts' activities and linkages in relation to T/E/I changed over time?

3. How do school districts currently manage linkages among T/E/I?

4. What impact do system-supported linkages have?

Our five interviewers are well trained and familiar with schools and districts; all have been teachers, two have been administrators. They are sensitive to the demands of the time of school personnel, as well as to the need for confidentiality and discretion.

We hope you share our enthusiasm for the proposed collaborative arrangement, and that we will have an opportunity to work together in the months ahead. We would appreciate your sending us a letter indicating whether your district will be able to join us and indicating the name, telephone number, and address of the person who we should call as the district's contact person.

Sincerely,

ORGANIZATION OF TRAINING FOR FIELDWORK

1. Orientation session - April 1, 1980

    A. Background of project, purposes and intentions

    B. Overview of fieldwork methodology

        need for consistency in question-asking

        need for common level of detail across interviews

        nature of evidence for conclusions

    C. Review of questionnaires for first visit

    D. Discussion of respondent sample

        common to all districts

        specific to situation

2. Reliability checking

    A. First round of interviews, in pairs, with subsequent debriefing and listening to tapes

    B. Listening to tapes and review of field notes

3. Retraining for reliability - before second visit

    A. Discussion of issues, problems

    B. Peer debriefing based on tapes and field notes

TRAINING MATERIALS
EVALUATION DESIGN PROJECT
April 1, 1980

INTERVIEW PROCEDURES

Before the interview begins

1. Know the intent of each question on the interview form.
2. Test tape recorder and tape.
3. Mark tape with name, date, place, etc., of interview.
4. Find a quiet place where there will be no interruptions.
5. Use standardized explanation of what we are doing, who we are, where you are from (see Attachment A).
6. Explain our position on confidentiality and uses of data (see Attachment B).
7. Explain likely length of interview, topics to be covered.
8. Request permission to use tape. If pushed, explain why needed.
9. Arrange recorder so it will pick up sound.
10. Complete face sheet.

During the interview

1. Stick as closely as possible to the wording of the interview questionnaire.
2. Ask follow-up probes as indicated on interview.
3. If respondent uses phrases or expressions you don't understand, ask: "What do you mean by ...?" "Can you give me an example of...?" "What would I see if I...?"
4. Stop and summarize if you don't understand. "My understanding is..." "You've told me so far... Is there anything else?"
5. Keep control of the interview. If respondent goes off on tangents, say: "I'd like to get back to..." "It would be best if we discussed that later, now let's talk about..." Use nonverbal cues (turning pages, restlessness).
6. Listen carefully. Stay concentrated. Don't assume you already know the answers.

7. Keep interview conversational rather than interrogatory. Be interested in responses. Don't enter into discussions. Don't give leading or evaluative feedback either verbally or with cues.

8. Indicate that interview information was very valuable to you. Thank respondent for time.

9. Take notes during interview as backup for tape and to indicate important things to listen for.

After interview ends

1. Take clean form and write up the interview while listening to the tape. (See Attachment C for instructions.)

2. Fill in record form (Attachment D).

3. Make entry in field journal (Attachment E). Use left-hand margin for interpretations, inferences, ideas. Use right-hand margin for anecdotes, incidents, etc., using as descriptive and low inference language as possible.

ATTACHMENT A

Let me start by telling you about the purpose of this interview. I'm _____ from UCLA. I work at a research institution in the Graduate School of Education, called the Center for the Study of Evaluation. The Center is one of 17 labs and centers funded by the National Institute of Education to do research and development in education. CSE has been doing this since 1965.

Currently, the Center is interested in two broad research areas - testing and evaluation. We've just begun a project to find out how school districts use testing or evaluation data to improve instruction. That's why we're here in the _____ district.

ATTACHMENT B

I want to assure you that our talk is confidential; the answers we get from you and other people in your school district will be combined into an overview. Where examples and quotations are used as evidence, people and places will not be identified.

So, the purpose of our interview is to help us understand the linkage between testing/evaluation and instruction. Any questions before we begin?

I'd like to get your permission to tape-record our talk, so I don't miss anything in my notes or change your words or ideas. Again, you won't be identified in the tape and only our project staff will hear it to make a transcription. If you want the recorder off at any time, just press this button. (Demonstrate)

ATTACHMENT C

WRITE UP THE INTERVIEW

Soon after completing the interview, you will go through the tape recording and write up the interview in detail. Spend no more than two hours per interview on this task. You'll need to use your judgment in deciding which pieces of the interview, what wording, to transcribe verbatim. (You are certainly not expected to transcribe the entire interview.)

In writing up your interview:

1. Use a new, blank interview form.

2. Feel free to continue answers on the backs of pages (but note the question number you're continuing).

3. Where you omit large pieces of talk, note this in parentheses, stating succinctly what has been omitted, e.g., (Here, the interviewee spoke about his experiences in another school, which he compared unfavorably with this one), or, (Here, the respondent spoke at length about the impact of her role on the SSC upon he preparing for class. She felt it took great amounts of time away from her daily planning.)

4. In general, try to quote exactly, but note where you are paraphrasing, rather than quoting: [Paraphrase]

5. If you asked a probe or rephrased a question such that you inquired into an area not on the interview form, note down the question you asked: (Question: ".....?")

6. Try to keep respondents' remarks in chronological order.

7. Keep the interview form with the notes you took during the interview.

POSSIBLE SPECIAL TOPICS/THEMES

Leadership

Coordination

Communication

Accountability

Role of principal

Morale

District continuum

Effect of community pressure

Effect of staff stability

Clarity of roles/responsibilities

Supervision

Teacher decisionmaking

Role of Learning Specialist

Factors being coordinated

School ethos

Influence of specific people

Problems of linking t/e/i

Impact of district actions on schools/classrooms

Effect of size of district

Centralization/decentralization

Differential use of tests

- Hart Bill
- CTBS
- CAP
- CRTs

Differential use of evaluations

- Title I
- Title 7
- SIP - PQR
- Other

Planning

Implementation

Phases